LUND UNIVERSITY

Application of automated video analysis to road user behaviour

Laureshyn, Aliaksei

2010

Link to publication

Citation for published version (APA):
Laureshyn, A. (2010). *Application of automated video analysis to road user behaviour*. [Doctoral Thesis (compilation), Transport and Roads].

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Application of automated video analysis to road user behaviour

Aliaksei Laureshyn

Doctoral Thesis CODEN:LUTVDG/(TVTT-1039)1-202/2010

Bulletin - Lunds Universitet, Tekniska högskolan i Lund, Institutionen för teknik och samhälle, 253

ISBN 978-91-628-8003-3
ISSN 1653-1930

Aliaksei Laureshyn

Application of automated video analysis to road user behaviour

2010

Keywords:
Video analysis, road user behaviour, traffic conflicts, traffic safety, safety indicators

Abstract:
The successful planning, design and management of a traffic system is impossible without knowledge of how the traffic environment affects the behaviour of road users and how the behaviour is related to the main qualities of the traffic system (e.g. safety, efficiency). Automated video analysis is a promising tool for traffic behaviour research in that it enables collection of micro-level behaviour data for large populations of road users and provides a detailed description of their motion. This thesis describes the tests done with an automated video analysis system developed at Lund University. The system was used in two large scale studies with the main task of detecting the presence of road users of a particular type. Accuracy of position and speed estimates were tested in a study specially designed for that purpose. The thesis also elaborates on the problem of relating the behaviour of road users to safety and proposes organising all the elementary events in traffic (defined here as encounters between two road users) into a severity hierarchy. The process of an encounter is described with a set of continuous safety indicators that can handle the various approach angles and transfer between being and not being on a collision course. When an objective measure for an encounter severity is found, the severity hierarchies may be used not only for describing safety but also for studying the balance between safety and other qualities valued by road users.

Citation:
Laureshyn, Aliaksei. Application of automated video analysis to road user behaviour. Institutionen för Teknik och samhälle, Trafik och väg, 2010. Bulletin - Lunds Universitet, Tekniska högskolan i Lund, Institutionen för teknik och samhälle, 253

Trafik och väg
Institutionen för Teknik och samhälle
Lunds Tekniska Högskola, LTH
Lunds universitet
Box 118, 221 00 Lund

Traffic and Roads
Department of Technology and Society
Faculty of Engineering, LTH
Lund University
Box 118, SE-221 00 Lund, Sweden

ISBN 978-91-628-8003-3
© Aliaksei Laureshyn, 2010
Printed in Sweden
Media-Tryck, Lund, 2010

CONTENTS

LIST OF PUBLICATIONS

GLOSSARY OF TERMS

1. INTRODUCTION 11
   1.1. Background 11
   1.2. Scope and objectives 14
   1.3. Thesis structure 14

2. INDICATORS IN TRAFFIC BEHAVIOUR STUDIES 17
   2.1. Quality of an indicator 17
   2.2. What indicators are used now – a literature study 18
   2.3. Indirect traffic safety indicators 19
   2.4. What is expected from a video analysis tool? 25

3. VIDEO ANALYSIS TECHNOLOGY 27
   3.1. What is video analysis? 27
   3.2. Traffic-oriented video analysis applications 28
   3.3. Video analysis system at Lund University 31
      3.3.1. Advanced road user detection 32
      3.3.2. Trajectory extraction I 33
      3.3.3. Rectification. 34
      3.3.4. Speed estimation 35
      3.3.5. Trajectory extraction II 35
      3.3.6. Interpretation of data extracted from video 36

4. STUDYING THE TRAFFIC PROCESS 39
   4.1. Continuous safety indicators 39
      4.1.1. Severity rating with time-based indicators 39
      4.1.2. Description of an encounter process with a set of indicators 45
   4.2. Behaviour classification by pattern recognition techniques 51
      4.2.1. From operational to tactical data 51
      4.2.2. Pattern recognition at work 53

5. PRACTICAL EXPERIENCE 57
   5.1. Study I – Cyclists on one-way streets in Stockholm 58
      5.1.1. Background 58
      5.1.2. Study design 58
      5.1.3. Results 59

5.2. Study II – Cyclists in roundabouts, 2 design solutions 61
5.2.1. Background 61
5.2.2. Study design 62
5.2.3. Results 63
5.3. Accuracy tests 64
5.4. Factors that affect the accuracy of the measurements taken from video 66
5.4.1. Video recording quality 66
5.4.2. Video processing algorithms 68
5.4.3. Traffic data interpretation algorithms 71
5.4.4. Accuracy and observation period 71

6. DISCUSSION 73
6.1. The video analysis system - current state and future development 73
6.2. Behavioural studies and safety evaluation based on video data 76

7. CONCLUSIONS 81

8. ACKNOWLEDGEMENT 83

9. REFERENCES 85

PAPER I

PAPER II

PAPER III

PAPER IV

PAPER V

LIST OF PUBLICATIONS

Paper I Laureshyn, A., Å. Svensson. "Road user behaviour indicators in automated video analysis systems". Submitted to *Journal of ITS* on 25th June 2008, re-submitted after a revision on 11th September 2009.

Paper II Laureshyn, A., Å. Svensson, C. Hydén. "Evaluation of traffic safety, based on micro-level behavioural data: theoretical framework and first implementation". Submitted to *Accident Analysis and Prevention* on 8th December 2009.

Paper III Laureshyn A., K. Åström, K. Brundell-Freij (2009) "From speed profile data to analysis of behaviour: classification by pattern recognition techniques". *IATSS Research* 33 (2), pp. 88-98.

Paper IV Laureshyn, A., H. Ardö, T. Jonsson, Å. Svensson (2009) "Application of automated video analysis for behavioural studies: concept and experience". *IET Intelligent Transport Systems* 3 (3), pp. 345-357.

Paper V Sakshaug, L., Å. Svensson, C. Hydén, A. Laureshyn. "Bicyclists in roundabouts – different design solutions". Submitted to *Accident Analysis and Prevention* on 26th September 2009.

GLOSSARY OF TERMS

Accident severity	some operational parameter describing the outcome of an accident
Collision course	a situation when the road users will collide if they continue with unchanged speeds and paths
Collision point	location of the first physical contact (projected on a road plane) when two road users collide
Conflicting Speed (CS)	in the Swedish Traffic Conflicts Technique: the speed of the *relevant road user* at the moment of the first evasive action taken by one of the road users
Crossing course	a situation when two road users pass a common spatial zone, but with some time margin and thus avoid a collision; for collision to become possible, a change in temporal relation of the road users is needed
Diverging course	a situation when the paths of two road users do not overlap and thus a collision is avoided; for collision to become possible, a correction in spatial and, possibly, a temporal relation of the road users is needed
Encounter	simultaneous presence of two road users within some pre-defined area
Encounter severity	an operational parameter describing the "closeness" of an *encounter* to a collision. Ideally, encounter severity should reflect both the risk of a collision and the severity of possible consequences
Indicator	objective and measurable parameter that has a relation to a studied quality of the traffic system (e.g. efficiency, safety, comfort, etc.
Near-miss	a situation when two road users unintentionally pass each other with a very small margin, so that the general feeling is that a collision was "near"
Relevant road user	in the Swedish Traffic Conflicts Technique: the road user that determines the severity of a traffic conflict

Reliability	the property of an *indicator* to be measured with the same accuracy and objectivity regardless to where, in what conditions and by whom the measurements are performed
Severity hierarchy	distribution of elementary events in traffic rated according to some operational severity measure
Speed profile	a continuous description of road user's speed over time; in a video analysis system a speed profile is represented as a sequence of speed measurements taken with high frequency
T_2	a complimentary parameter to *Time Advantage* describing the time for the second road user to arrive at the *collision point*
Time Advantage (TAdv)	a minimal correction in time (a delay of one of the road users) that is necessary for road users to come on a *collision course*
Time Gap (TG)	a parameter describing the spatial proximity of two road users expressed in time units
Time-to-Accident (TA)	in the Swedish Traffic Conflicts Technique: the time remaining from the first evasive action taken by one of the road users up to the collision that might have taken place had they continued with unchanged speeds and paths
Time-to-Collision (TTC)	in *collision-course* situation: the time required for two vehicles to collide if they continue at their present speeds and on the same paths
Traffic conflict	an observable situation in which two or more road users approach each other in space and time to such an extent that a collision is imminent if their movements remain unchanged
Traffic-conflict technique	a method for traffic safety estimation based on observation of *traffic conflicts*. The basic hypothesis of traffic-conflict techniques is that accidents and conflicts originate from the same type of processes in traffic and a relation between them can be found
Trajectory	a path of a road user on the road plane; in a video analysis system a trajectory is represented as a sequence of positions measured with high frequency
Validity	the property of an *indicator* to describe the quality that it is intended to represent

1. INTRODUCTION

1.1. Background

The successful planning, design and management of a traffic system is impossible without knowledge of how the traffic environment affects the behaviour of road users and what the relation is between the behaviour and main qualities of the traffic system, i.e., its safety, efficiency and comfort. Answering these questions helps to understand what behaviour is desired and how it can be promoted.

The behaviour of an individual road user is the basic data unit in a bottom-up data collection. For example, the total emission at an intersection is made up of the emissions of each vehicle, which in turn depend on the vehicle type and type of regulation of the intersection, but also very much on how the vehicle passed the intersection, i.e., if it stopped, waited, inched forward or drove without changes in speed. Individual interactions can say a lot about the functioning of the intersection, providing data for calculation of aggregated parameters like the average number of stops, delays, traffic work and exposure. They can also indicate possible problems, for instance misunderstanding of the design or high complexity of certain situations that lead to road users' distraction and possible errors. Studies of the breakdowns in the interaction process that result in situations close to a collision (traffic conflicts) have a great potential for safety estimation. A considerable advantage of such studies is that the accident potential and the processes leading to accidents may be judged without observing the actual accidents.

Generally, behavioural studies differ from other approaches in psychology in that many abstract concepts like "personality", "attitude", "motivation", etc. are avoided and efforts are concentrated on collecting and analysing data on objective actions – the outcome of all the internal psychological processes. Applied to traffic, a behavioural study means examination of measurable parameters like speed, position, distance, observable signals and actions, etc. and their relation to conditions and factors in the road environment and actions of other road users.

Collecting data on behaviour in traffic is not a simple task. Although the list of parameters (indicators) that describe the behaviour can be very long, many of them are difficult to measure with conventional instruments like radar guns or inductive loops. So far, using human observers has been the most common practice, but there are many limitations related to this approach. These include decreasing attention as the observer gets tired, high costs (which seriously limit the length of the observation time), risk of subjective judgements, possible effects on road users' behaviour when

they find themselves being observed and problems with finding a position for the observer so that his/her vision is not obscured. In addition, there is always a risk of the observer being involved in or causing an accident while making observations in the middle of a complex traffic environment.

A video recording has many advantages compared to the road-side observations and helps to avoid some of these problems. A camera is more discrete and not so easily detected by road users, recordings can be run autonomously for longer periods and analysed indoors in more comfortable conditions later, and there is an opportunity to look at the relevant situations again and study them in detail. Still, some difficulties associated with human observations as a detection method remain. It is still the observer who makes the necessary measurements and detects the occurrences of interest. Watching a video film often takes the same time as on-site observations. At the same time, the observer's perception of the scene under study is more restricted since video is only a flat representation of the 3-dimensional reality. It is not possible to "turn the head" to follow a particular road user for a longer time or use the surrounding sounds to guide attention in the same way as it is done on-site.

Building a tool that will facilitate the analysis of video recordings is a logical development of the method. The range of techniques is wide – from simple software solutions that help to browse through video files and take some measurements by mouse-clicking, to very advanced systems that can automatically detect and follow road users and even measure behavioural indicators or detect special situations that are relevant for the study. Automation makes the data collection more efficient and systematic and makes it possible, in many cases, to skip watching the entire recorded video and concentrate only on the relevant parts of it. It contributes also to standardisation of the methods for behaviour data collection and analysis.

It appears, though, that the attempts to apply advanced video analysis techniques for advanced traffic behaviour research are quite rare at the moment. A possible explanation might be that behavioural studies are often done in complex traffic conditions with a great variety of possible trajectories and road users of different types (e.g. vehicles, pedestrians, cyclists, etc.) mixed together. This complicates the detection task, while the need for detailed description of the road users' actions puts high requirements on the accuracy of the extracted position and other types of data. The available video analysis technology is just starting to be mature enough to meet these requirements. Another possible reason is that the communication between traffic behaviour researchers and developers of the video analysis systems is not yet properly established – the former know too little about the available techniques, their advantages and weaknesses, while the latter need more detailed explanations of what qualities are crucial for the technology to be suitable for behavioural research.

This thesis describes the work done in a research project at the Faculty of Engineering, LTH, Lund University, which took place during 2004-2008, with the aim of developing a prototype of an automated video analysis system with traffic behaviour studies as a primary application area. Two main actors were involved, representing two research traditions that met in the project (see Figure 1). The Department of Technology and Society, Traffic and Roads, was the initiator of the

project and formulated the traffic-related questions for which the video analysis system could be useful. It was responsible for the tasks of interpretation of the road users' behaviour in traffic terms, development of the theories for how behaviour is related to the traffic system qualities, etc. The Centre for Mathematical Sciences was responsible for the tasks related to the development of video processing algorithms, such as detection and tracking of road users and transformation from image to road plane co-ordinates. Some tasks, marked as a "grey area" in Figure 1, did not belong to the traditional domain of either of the actors, but were highly important for the successful project work. These were, for example, problems related to camera installations, modelling of the typical road users' shapes, organisation and storage of the data, accuracy tests, etc. Both actors had to extend their efforts outside the "traditional" research areas in order to have the "grey" area covered and elaborated, too.

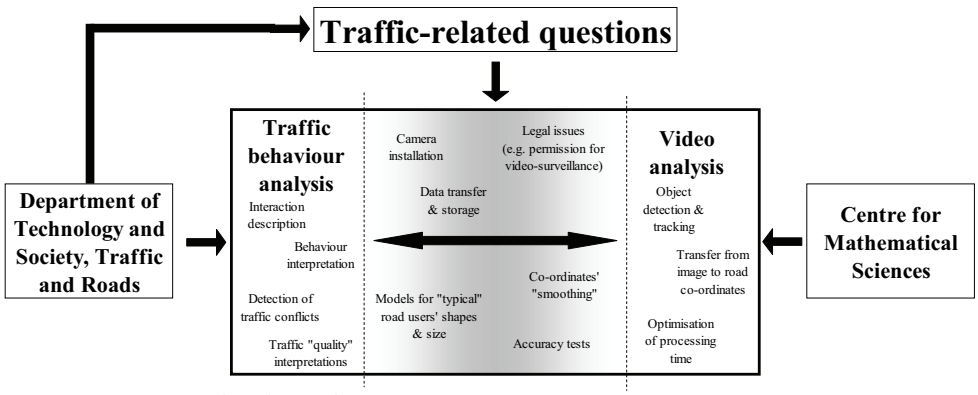


Figure 1. Actors and tasks in the project

The interest and motivation for the project came through several earlier attempts to make traffic behaviour data collection more effective by using the video analysis solutions available at that time. The efforts were largely concentrated on improving reliability and making the Swedish Traffic Conflicts Technique (*Hydén, 1987*) more usable by automation of the conflict detection procedures. Tests with semi-automated video analysis systems (*Andersson, 2000, Odelid et al., 1991*) showed the great potential of using video data, but the work on manual clicking of the road users was too tedious and time-demanding in a large-scale study and the need for automation was obvious. However, the experience with existing automated systems was not very successful (*Odelid & Svensson, 1993*). It became clear very early on that the detection of conflicts is an ultimate task that requires performance much higher than shown by the tested solutions, and that some important issues that seemed obvious for a traffic researcher were not prioritised at all by the producers of the video analysis tools (e.g. that position of a road user should be estimated in road plane co-ordinates and not only in image co-ordinates). On the other hand, these tests stimulated thoughts about many other possible applications of video analysis, where the tasks are not as complex as in conflict detection and therefore even simpler technologies may be used if properly adjusted. It is necessary to systematically map the expectations from a data collection tool and the functionality of the automated video analysis technology, to

reveal where these two sides meet and do not meet and to test the application of the video analysis in traffic research on a larger scale.

1.2. Scope and objectives

This thesis is a part of the ongoing work on development of a video analysis system at Lund University to facilitate behavioural studies in traffic. From the very beginning, this development was governed by the needs of traffic research, and not the other way round ("we have a technology, so where can we use it?"). The work in practice consists of constant discussion and exchange of ideas, requirements and experience between the traffic and video analysis parts of the team. The system is based on the following principles:

- The system uses video recordings obtained from a camera/cameras statically fixed in the road environment at a significant height;
- The road users are detected in the recorded video and their trajectories are produced;
- The ultimate goal is to make video processing as close to real time as possible. While this has not been achieved yet, the offline processing is accepted as a sufficient temporary solution.

The objectives for this thesis are as follows:

1. To explore the current practice in traffic behavioural research and estimate the potential of using a video analysis system in this area.
2. To investigate the type and quality of the results produced by video analysis technology and how they may be used for behavioural studies.
3. To suggest and test more advanced indicators and analysis methods that can be applied on data produced by video analysis.
4. To test video analysis technology in large-scale behavioural studies, present the results and discuss the lessons learned.
5. To contribute to better communication and understanding between the developers of the video processing algorithm and the traffic researchers.

1.3. Thesis structure

The thesis has the following structure. After the Introduction (Chapter 1), I discuss what indicators are used in traffic behaviour studies, especially the indirect safety indicators, and which of them can be collected from video data. I formulate also the expectations from an "ideal" video analysis system that can be used in traffic behaviour research (Chapter 2). Chapter 3 provides a brief overview of the principles of video analysis technology, and describes the system developed within the project and used in the tests. Chapter 4 elaborates on the analysis of the sequential data extracted from video. Chapter 5 describes the studies where video analysis was used as a main measurement tool and discusses the factors that affect the accuracy of the

measurements taken from video. Chapter 6 presents a final discussion followed by Conclusions (Chapter 7). The thesis also includes five articles published or submitted for publishing in scientific journals. The articles elaborate in more detail some of the topics of the thesis.

The detailed technical information on the video analysis techniques used in this research can be found in *Ardö, 2009*, which is another doctoral thesis produced within the same project.

2. INDICATORS IN TRAFFIC BEHAVIOUR STUDIES

2.1. Quality of an indicator

As there is a great variety of road users' actions that might be relevant in a behavioural study, a certain degree of generalization and simplification is necessary in order to make the study practically feasible. Usually, we are restricted to collecting data on just a few indicators, i.e., objective and measurable parameters that we believe have a relation to the studied qualities of the traffic system (efficiency, safety, comfort, accessibility, etc.)

The two key properties of an indicator are its ***validity*** and ***reliability***. The validity refers to whether an indicator describes the quality that it is intended to represent and to what extent. For example, red-walking at a pedestrian crossing is certainly related to the safety of pedestrians. However, crossing on red after having waited for a long time with no on-coming vehicles present is probably not very risky, while a pedestrian who arrives in a hurry and continues over the crossing without looking around puts himself in an extremely unsafe situation. The frequency of the second type of red-walking is presumably a more valid indicator compared to that of all red-walkers without distinguishing them in any way.

Establishing the validity of an indicator usually requires numerous large-scale studies performed in various conditions. The correlations found have to be supported by a theory providing clear logical and causal connection between the indicator and the quality it is supposed to represent, and possible confounding factors have to be controlled for (*Elvik, 2008*).

Reliability refers to the methods used to measure the indicator and the accuracy of the measurements. The accuracy of a reliable method should remain within the same limits regardless of measurement locations, time of the day and traffic conditions, thus ensuring that the difference in the results reflects the difference in the studied phenomenon and not in the measurement's accuracy. For example, much criticism has been directed towards the traffic-conflict techniques' complete reliance on a human observer's subjective judgements of the distances between road users and their speeds. However, the tests comparing the estimations of different observers and the objective measurements taken from a video film showed that the results were very similar, i.e., the method was proven to be quite reliable (*Hydén, 1987, Asmussen, 1984*).

A starting point in assessing how video analysis is applicable to traffic behaviour studies is to have a look at the current practice, especially what indicators are typically collected and if the same indicators can be measured using video analysis. Our department has traditionally had a special interest in traffic safety studies and safety indicators, therefore this topic is further discussed in a separate section.

2.2. What indicators are used now – a literature study

To find out what indicators are commonly used in behavioural studies and which of them, theoretically, can be retrieved from video data, I have carried out a literature survey and made up a snapshot list of indicators currently in use, classified into groups according to the type of behaviour and types of road users they represent.

The study covers 45 relatively recent research articles and reports. The main criterion for a study to be selected for this review is its contribution of new indicators to the list; therefore neither the quality of the study design nor the results have been judged strictly. However, only articles from reviewed scientific journals, doctoral theses and reports from well-established research institutes are included, for reasons of credibility.

For each indicator, the following information is retrieved:

- Type of the road user involved – vehicle drivers, pedestrians and cyclists; I have treated indicators describing traffic conflicts as a separate group, as, firstly, they are a very special type of traffic situation and, secondly, all the conflict indicators are universal and may be applied to any type of a road user;
- The type of property described by the indicator – an individual road user's properties (like age, gender, etc.), individual behaviour, interaction with other road users or the property of the traffic environment on the aggregated level;
- The type of data provided by an indicator – binary ("yes/no", e.g. if a car stops or not), a single value (e.g. pedestrian's age) or a sequence of values (for example, a trajectory, i.e., a sequence of positions over time).
- If an indicator can be derived from the trajectory co-ordinates of the road users or not, i.e., if it can be calculated from the data produced by video processing algorithms.

The detailed description of the reviewed indicators and the summarising indicator lists can be found in Paper I.

Totally, the literature study has yielded 119 unique indicators. The review suggests that 98 of them (i.e. 86%) can be expressed through road users' co-ordinates and parameters like speed, direction, etc. (i.e., those calculated from the trajectories data). Some of them, however, require additional input from other measuring instruments; for example, to decide if a vehicle arrives on green, yellow or red, simultaneous information from the traffic light is also necessary. The remaining indicators describe the personal characteristics of road users (e.g. age and sex), and actions like head, eye and hand movements and eye contact.

The review also shows that indicators of the "yes/no"- and “single value”-types are dominating. Some of the indicators can only be described in this way, for example, the question of whether a pedestrian crosses a street before or after a car. Many other indicators may easily be modified so that they represent a sequence of values when an instrument capable of measuring the parameter with high frequency is available (e.g. vehicle speed can be measured at a certain fixed point, or as a speed profile over time).

Many of the reviewed indicators are used in traffic safety context, though their validity as safety indicators is not always unquestionable. I elaborate further on the indirect safety measures in the next section, where I make an overview of the concept and state the problematic areas that have to be resolved to ensure the indicators' validity. Since most of the indicators can be retrieved from video data, contribution of the video analysis to that as a data collection tool is very important.

2.3. Indirect traffic safety indicators

For many traffic safety studies the accident history is the main, and often the only, input data source. While accidents are an obvious safety indicator, accident analysis as a safety research method has some quite important limitations that have been intensively discussed in the literature (e.g. *Elvik & Vaa, 2004, Berntman, 2003, Englund et al., 1998*). Some of the main concerns about accident analysis are:

- Accidents are rare events and it takes a long time to collect a sufficient amount of accident data to produce reliable estimates of traffic safety, e.g. the expected number of accidents per year. For longer analysed periods it is hard to associate the change in accident number with a particular factor as the other relevant factors might also change during this time. There is also an ethical problem in that one has to wait for sufficient number of accidents to occur and thus for people to suffer before anything can be said about safety;
- Accidents are random and the number of accidents registered every year at the same place is not the same, even if the traffic situation does not change. If one year has an unusually high number of accidents, one could expect that the number of accidents will go down in the next year, which is just the natural fluctuation around some “average” accident level. This phenomenon is known as the “regression-to-mean” and is dealt with in many connections (e.g. *Elvik & Vaa, 2004, Hauer, 1997*). From this perspective, the actual accident number is also an indirect measure, while the “true” safety characteristic is the “expected number of accidents” that cannot be measured but only estimated based on the accident history or other safety indicators (*Hauer, 1997*).
- Not all accidents are reported. The level of underreporting depends on the accident severity and types of road users involved. In Sweden, for instance, the level of reporting for fatal and severe injury accidents is near 100%, while slight injury accidents and especially property-damage-only accidents are reported very poorly. Comparison of police accident data with hospital admission records reveals that vulnerable road users (pedestrian, cyclists) are greatly

underrepresented in accident statistics. For this reason, it is desirable that accident databases include the data from both police and hospital records (as implemented, for example, in STRADA, the Swedish national accident database – *Transportstyrelsen, 2009*).

- Important information on the process of accident development (linking the behaviour of the road users to the accident) is often missing in accident reports. Reconstruction of an accident to collect such data is usually very costly and not always possible to perform; doing it with a help of witnesses and those involved in the accident presents a great risk of bias. Without understanding the process that leads to an accident and the role of the infrastructure design in it, it is very hard to propose measures to change the behaviour that results in accidents.

For this reason, there is a growing interest in using some other, indirect measures for traffic safety (also called surrogate safety measures or safety performance indicators - *Gettman & Head, 2003, ETSC, 2001*). Indirect in this context means that these measures are not based on accidents, but rather on other occurrences in traffic that are "… causally related to crashes or injuries, … <and> used in order to indicate safety performance or understand the process that leads to accidents" (*ETSC, 2001*). The goal is to estimate the "expected number of accidents", not to predict the actual accident occurrences. In fact, the indirect indicators might be as good as the actual accidents in estimating the expected number of accidents (*Svensson, 1992*).

In an ideal case, an indirect safety measure should fulfil the following criteria:

- To be relatively frequently observed in traffic. This shortens the time necessary for data collection and allows quick safety evaluation studies to be made before the actual accidents occur.
- The methods for indicator data collection should be objective and reliable.
- To have clear logical and strong statistical relations to accidents. This is an important condition for the validity of a safety indicator.
- There should be some degree of similarity between the situations described by a safety indicator and accidents, which allows us to study the process of accident development without an actual crash or injury.

Many indirect safety indicators have been proposed. For example, the European Traffic Safety Council recommends regular monitoring of road user-related indicators like speeding, drunk driving and use of restraint systems and safety devices (*ETSC, 2001*). The interim targets for road safety in Sweden, proposed by the Swedish Road Administration for the period until 2020, include control of 13 indirect safety indicators, of which many are road user-related, e.g. speeding, drunk driving, use of helmets and seat belts, and drivers' fatigue (*Vägverket, 2008*).

An important distinction may be drawn between these indicators; some of them reflect the probability of an accident (e.g. speeding or drunk driving), while others reflect the probability of fatalities and injuries *if* an accident occurs (e.g. use of safety belts, helmets). The validity of the latter type is relatively easy to establish as the functioning mechanisms of a protection device are obvious and it is even possible to

test its efficiency in a laboratory crash-test. This is not the case for the indicators reflecting the probability of an accident and so far very few indicators have been thoroughly validated. An increase in general speed level and speeding by individual road users is proven to be strongly associated with both higher accident risks and the severity of consequences (*Elvik et al., 2004, Nilsson, 2004, Quimby et al., 1999, Maycock et al., 1998, Kloeden et al., 1997, Fildes et al., 1991*). The high risk of alcohol-influenced driving is also an established fact (*Arranz & Gil, 2009, GRSP, 2007, Evans, 1991*). As for fatigue, researchers appear to be in agreement about its importance as a contributing risk factor, but the absence of a tool for an objective fatigue measurement (e.g. breath analysers for drunk-driver control) complicates the introduction of usable indicators and their validation (*Radun & Radun, 2009*).

The described indicators are mostly meant for monitoring purposes and testing specific hypothesis regarding safety, not for explaining the behaviour leading to accidents. It is even possible that some of them are not very good descriptors of safety (for example, use of cycle helmets does not automatically increase safety as it may result in more risky behaviour of cyclists that compensate the positive effect of helmets use). Still, the information they provide is important to get the general overview of the state of the traffic system and indicate that some safety measures may not work as well as expected.

A special category of indirect safety measures (maybe the most "direct" of them) that is particularly concerned with the individual behaviour and the process of interaction between road users is traffic conflicts. A traffic conflict was defined at the first international workshop on traffic-conflict techniques in 1977 as "an observable situation in which two or more road users approach each other in space and time to such an extent that a collision is imminent if their movements remain unchanged" (*Amundsen & Hydén, 1977*, cited *Hydén, 1987*). According to this definition, an accident is always preceded by a traffic conflict, but most often conflicts do not develop to such a degree that a collision occurs, as at least one of the road users takes some evasive action. Still, the conflicts can say a lot about the accidents as the underlying processes are very similar.

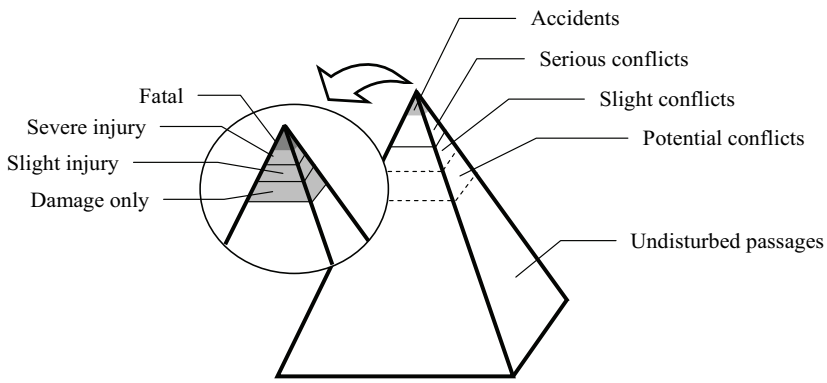


Figure 2. The pyramid-model (**Hydén, 1987**), describing the relation between severity of the elementary events in traffic and their frequency.

Hydén, 1987, proposes a general model describing the relation between normal events in traffic, traffic conflicts and accidents, illustrated with a pyramid (Figure 2). The top of the pyramid represents accidents which are the most severe but also the most rare events in traffic. It is also possible to distinguish fatal, injury and damage only accidents, and as the severity of accidents goes down they become more frequent. Just below the accidents in the pyramid come traffic conflicts ("near-accidents") that may also be classified as serious, slight or potential conflicts according to their severity. Below the conflicts come the majority of the events that characterise the normal traffic process. This model suggests that there is some severity dimension common to all the elementary events, defined in general terms as the closeness to an accident and severity of its consequences.

The traffic process consists of interactions between road users. A simplest "unit" of interaction is an encounter between two road users (if there are more than two road users involved, the situation can still be considered as a combination of several pair-wise encounters). It seems reasonable to assume the existence of a relation, similar to Hydén's pyramid, between the frequency of the encounters and their severity level.

The severity of an encounter without a collision should reflect both the potential of an encounter to become an accident *and* the possible consequences. The former may be explained in the following way. Accidents are stochastic events. Even though one particular accident may be explained by a number of factors that led to it (road and vehicle conditions, driver's emotional and physical state, traffic situation, possible disturbance by a third road user, etc.), it can be considered as an unlucky coincidence that all these factors happened to be present at the same time. If some of these factors had not been present, the accident might have been avoided. Put in another way, each encounter between two road users has the potential to develop into an accident if, by a coincidence, some additional unlucky factors appear. However, a near-miss has less of a safety margin to endure an additional unlucky factor compared to a well-controlled passage; thus the severity of a near-miss is higher.

The second aspect that a severity of an encounter has to reflect is the consequences *if* an accident occurs. Since we deal with a hypothetical accident and its consequences are not determined, there is a set of probabilities that describe the potential of an encounter to develop into an accident of certain severity.

Having a severity measure for each encounter, the encounters can be placed in some distribution as a function of severity (this kind of distributions is called severity hierarchy in *Svensson, 1998*). The way the severity is determined defines the actual shape of the hierarchy. It is reasonable to assume that there is a "true" hierarchy which reflects the objective severities of all the encounters. By introducing various operational measures to describe the severity, one may come up with many quite different hierarchies, in which the same event probably will not be placed exactly on the same level.

The severity hierarchy of the encounters at a particular site (e.g. an intersection) is more likely to be peaked at both the most severe and least severe ends. The least severe encounters are also quite rare, while the majority of the encounters are of "medium severity". Road users tend to "optimise" their travel by not keeping too

large safety margins since they loose in efficiency, but not making them too small either since it compromises safety. *Svensson, 1998*, also argues for the diamond shape of the distribution, but she limits the events included in the hierarchy to only those with a collision course (i.e. at some point the road users will collide if they continue with unchanged speeds and paths).

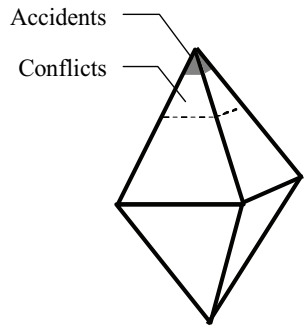


Figure 3. The diamond shape of a severity hierarchy (adopted from *Svensson, 1998)***.**

The frequency of events in the different severity levels of the hierarchy bears different information. Serious conflicts (according to the operational definition used in the Swedish Traffic Conflicts Technique) come just after accidents in the severity hierarchy and are very similar to the actual accidents. One of the strong arguments for this was found during the interviews with road users just after they were involved in serious conflicts. Most of them had a strong emotional reaction after the incident and stated that they did not want to appear in such a situation deliberately. This was not the case with non-serious conflicts where many interviewees could not remember the incident at all. A strong statistical correlation between the number of serious conflicts and the number of police reported injuries was also found. Moreover, when the severity for reported accidents was defined in the same terms as for the conflicts so that they could be analysed together, there was a clear pattern that more and more conflicts resulted in accidents as the severity increased with some "grey" zone in the middle with both accidents and serious conflicts that did not result in a collision present. This supported the hypothesis that serious conflicts and accidents originate from the same processes in traffic and thus development of accidents can be studied through analysis of development of conflicts (*Hydén, 1987*).

What information is given by the frequency of the encounters at the lower severity levels is not that clear. One hypothesis may be that a high number of non-serious conflicts also indicates unsafety. An alternative hypothesis may be that such situations do not result in many accidents, but they are "instructive" for the road users and make them more cautious and skilled, i.e. they are positive from a safety perspective. Thus, a large number of near-serious conflicts, while the serious conflicts are few, may indicate good performance of the traffic system. Findings in *Svensson, 1998*, support the latter hypothesis.

The best way to define the severity of an encounter is still an open issue. Various measures may be relevant, for example, proximity in space, proximity in time, intensity of a necessary evasive action, etc. These measures largely reflect the

probability of a collision. The outcome of the collision if the accident occurs depends on the kinetic energy released during the crash (i.e. mass and speed of road users), collision angle, types and vulnerability of the involved road users, etc.

The traffic-conflict techniques developed in different countries base the operational definitions of conflict severity on quite different parameters, from merely verbal descriptions for each category (*Baguley, 1984*) to using time- and space-related parameters requiring quite exact measurements (*Hydén, 1987, van der Horst & Kraay, 1986, Cooper, 1984*). However, good correspondence of the different techniques was found in the international calibration study in Malmö in 1984. The classification of conflicts by different techniques seemed to reflect some dimension common for all the techniques (*Grayson, 1984*).

The subjective perception of the situation by road users and observers most probably includes both the probability of a collision and its possible consequences. It is hard to deny that, being a pedestrian, a meeting with a large vehicle is experienced as much more risky than a meeting with a cyclist approaching with the same speed and at the same distance. In this respect, the subjective perception comes closer to the "true" severity compared to many objective measures that most often reflect just one of the severity aspects. This is supported by results of a validation study of the Swedish Traffic Conflicts Technique (*Svensson, 1992*), where the conflicts classified as serious based on subjective observers' judgments correlated better with the reported accidents than the serious conflicts classified strictly after the objective definition.

The perceived risk is a key issue in the theories dealing with road users' risk-taking behaviour in traffic. For example, the risk homeostasis theory (*Wilde, 1994*, see Figure 4) states that a road user attempts to keep a target level of risk based on some subjective estimates of costs and benefits of taking a risky action. This target risk is constantly compared to the current perceived level of risk, which derives from the direct perception of the situation, through a lagged feedback from past experience and general information about accident risks obtained from the media, discussions with colleagues and friends, official statistical reports, etc.

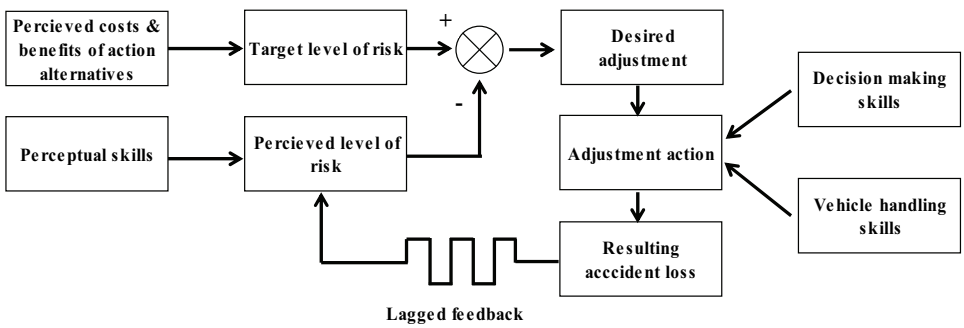


Figure 4. Homeostatic model relating the accident rate to the level of caution in road user behaviour (*Wilde, 1994****).***

Even though the risk homeostasis theory has received a lot of critic (e.g. *Evans, 1986*), the phenomenon of behaviour adaptation and compensation is generally recognised and accepted (*Englund et al., 1998*). One of the most serious complications occurs

when the perceived risk does not reflect the actual objective risk of the situation, causing behaviour modifications. A well-known example is the high number of accidents at zebra-crossings with priority given to pedestrians. A theoretical explanation is that pedestrians experience such crossings as very safe and pay very little attention to traffic, while the drivers sometimes fail to see the pedestrian or simply neglect the rules in order to maintain speed, feel power over a pedestrian, etc. (*Ekman, 1996*). Without knowing the mechanism of the subjective risk estimation by pedestrians, it is hard to propose measures that can remove the discrepancy between the subjective and objective risks.

The entire shape of the severity hierarchy can be seen as an illustration of a trade-off, made by road users, between their perceived risks and other (perceived) qualities like efficiency, mobility, comfort, etc. Again, it may be hypothesised that there is a "perfect" hierarchy shape where these qualities are optimally balanced. The deviations from the "perfect" hierarchy shape might indicate the discrepancies between the perceived and actual risks.

2.4. What is expected from a video analysis tool?

The fact that most of the indicators used in traffic behaviour studies may be retrieved from video data is very encouraging. However, there are many other factors that determine the usability of video analysis technology as a measurement and data collection tool. Keeping in mind the possible applications of such a tool, a traffic researcher's expectations of an "ideal" video analysis system may be stated as:

- A video analysis system has to provide detailed descriptions of road users' movements (in terms of road-plane co-ordinates related to time) with high space and temporal resolution to allow calculation of various (and complex) behaviour indicators and study the *processes* in traffic.
- It has to cover an area large enough to be considered as a logical unit of a traffic infrastructure, e.g. a pedestrian crossing, an intersection, a road section, etc., including the surroundings from where the behaviour may be influenced.
- It has to be effective enough to allow studied periods to be extended to the order of months and, possibly, years to be able to collect information about the most rare events in traffic (serious conflicts and accidents). The efficiency problem concerns, primarily, the time necessary for the video processing, but also the routines and technical solutions for video data storage, interpretation of the video processing results in traffic terms, effective presentation of the results, etc.
- The system has to provide a sufficient accuracy of road-user detection and tracking, their classification by type, position and speed measurements.

Hopefully, these expectations can make it easier for the developers of video analysis systems to understand what functionality is still missing and what problems require extra attention, as there are bottlenecks limiting the wider use of video analysis in traffic research. A full-fledged discussion, however, is not possible without both parties having an idea about their specific problems. Therefore, the next chapter

provides a short overview of the existing video analysis technologies and their traffic applications; it also describes the principles of the video analysis system developed at Lund University.

3. VIDEO ANALYSIS TECHNOLOGY

3.1. What is video analysis?

A digital video recording of a traffic environment contains a lot of information. One can see where the cars and pedestrians are moving, where tree branches are swaying or birds are flying, or just empty asphalt where nothing is happening at all. A human mind has no problem distinguishing traffic-related information from everything else and concentrating on it. However, programming a computer to do the same is no trivial task. A digital video consists of many still images taken one by one, each in its turn consisting of lots of pixels, i.e., small units that are assumed to have the same colour and intensity. The raw video data is thus just a set of values describing each pixel with a certain time frequency. With this enormous amount of data, it is a challenge to find a description of road users and transfer it into more usable traffic-related data like trajectories, speeds, size and orientation, etc.

The processing of a digital video usually combines two types of algorithms - analysis of the structure of each image and analysis of the changes in a sequence of images. Some pre-knowledge about road users is very helpful here. First of all, road users move, i.e., they come into the scene, pass through it (not necessarily without stopping) and then leave it. The detection of those not moving for a very long time (e.g. parked vehicles) is often not very important and such objects can simply be considered as a part of the static road environment (however, even though they do not take part in any interactions, they might affect the way the other road users interact). Despite the variety in the shapes of vehicles and constantly changing shapes of cyclists and pedestrians, we may still have an idea of the size of the objects we are dealing with. Very often, the size is also a good indication of the type of a detected road user. Vehicles appear most frequently on roads and pedestrians on sidewalks or pedestrian crossings. They always move on the ground surface which most often can be approximated as a plane.

There are also several well-known problems that complicate the detection and tracking of road users:

- Changes in the ambient light make even a static environment look different and may provoke false detections (e.g. an approaching cloud shadow is detected as an object moving on asphalt).

- The shadows of road users follow their paths and, if not removed, make the road users look as if they are taking more space in the image than they are. This produces errors in position and size estimations. Shadows may also complicate the separation of individual road users as they can "glue" several of them into one large object.
- Seen from the side, road users can occlude each other when coming close together. This complicates the separation of individual road users, leads to possible detection misses (if a smaller road user is not seen behind a larger one) and "swapping" of the identities of the road users when tracking after the occlusion.
- The road surface is rarely seen in dense traffic. Hence, the background model, describing the "typical" colour of the road when free from road users, becomes less reliable.
- An image is a 2-dimensional projection of a 3-dimensional reality and some information important for restoration of the true shapes of the road users (to estimate their true position) is missing.

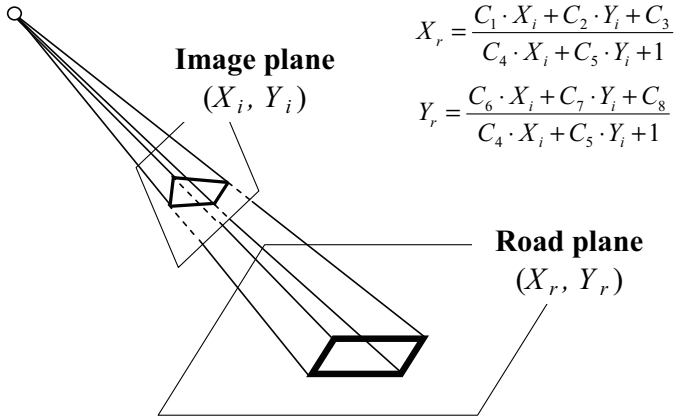
The extracted trajectories and other data usually require some additional interpretation in traffic terms, for example, finding road users performing a certain type of manoeuvre or detection of special situations like interactions, traffic conflicts, abnormal behaviour, etc.

Video analysis implies managing large volumes of data and intensive calculations, which also means that it takes a considerable time to obtain the results. This limits the lengths of the video recordings that can be processed within a reasonable time. An important factor here is the parameters of the hardware used for computations. However, the progress in this field during the last decades raises some hopes that this limitation will become less important in the future. The prediction of G. Moore done in 1960s that the performance of microprocessor chips will double about every two years (generally known as Moore's law, *Moore, 1965*) describes quite accurately the development of the semiconductor technology until now and it seems to remain true in the foreseeable future (*Thompson & Parthasarathy, 2006*).

3.2. Traffic-oriented video analysis applications

The first attempts to employ computer aid in processing video records were semi-automated systems. For example, a German system VIVAtraffic (*Rudolph, 1996*) provides a user-friendly interface for navigation through the video file frame by frame, and it is possible for an operator to take interactive measurements of road users' position and speed, as well as distances between a road user and other objects by "clicking" on them with a mouse. The points chosen for "clicking" have to be on a road plane (e.g. a point of contact between the vehicle wheel and the road or pedestrian's feet position), which makes it possible to change the "image" co-ordinates into "road" co-ordinates with a very simple transformation function (Figure 5).

Other semi-automated systems like VIDARTS (*van der Horst, 1990*), Trajex (*Andersson, 2000*) and SAVA (*Archer, 2005*), even though some of them were developed later and using more modern programming components, do not offer much additional functionality. The main problem of such systems is the very intensive manual work necessary for data extraction (for example, *Archer, 2005*, mentions that it takes 10 hours to process 1 hour of the original video). On the other hand, the collected data has high quality and is very relevant, as the operator works as an initial filter and selects only situations that are of interest. Despite the drawbacks, it is not unusual to see semi-automated systems being used even in relatively recent research (*Aronsson, 2005, Räsänen, 2005*).



$$X_r = \frac{C_1 \cdot X_i + C_2 \cdot Y_i + C_3}{C_4 \cdot X_i + C_5 \cdot Y_i + 1}$$

$$Y_r = \frac{C_6 \cdot X_i + C_7 \cdot Y_i + C_8}{C_4 \cdot X_i + C_5 \cdot Y_i + 1}$$

Figure 5. Transformation of the image co-ordinates into road-plane co-ordinates. Theoretically, knowing co-ordinates of 4 points on both image and road planes (given that any three of the points are not on the same line) allows us to find the transfer function between image plane (X_i , Y_i) and road plane (X_r , Y_r) co-ordinates (C_1 - C_8 are constants).

The simplest fully-automated systems are designed for tasks like the detection of presence and are based on the concept of a "virtual loop" (*Yung & Lai, 2001*), i.e., a certain manually defined area in the image monitored for changes in colour/intensity. When a vehicle passes this area, it looks quite different from the normal road appearance, and if the difference is above a certain threshold level, the detection is triggered. Several such "loops" may be monitored at a time. These algorithms are relatively simple, fast, and can even be run on a small processor embedded in a camera, which makes them very attractive as an alternative to conventional vehicle detectors. However, the application area is limited to conditions where the possible vehicles' positions are clearly defined and with no interference like other road users passing the same area but in another direction. A range of commercial products utilise this approach (*Autoscope®, 2009, Hitachi, 2009, Traficon®, 2008, VisioWay, 2007*).

The tracking of vehicles is successfully performed in simple traffic conditions, for example, on motorways. These systems use various approaches like feature tracking, shape models, data association, Markov chain, Monte Carlo simulation, wire-frame

models and line segment matching (*Schoenemann & Cremers, 2008, Zhang et al., 2008, Song & Nevatia, 2007, Coifman et al., 1998, Tan et al., 1998, Koller et al., 1992*). There are also classical solutions that can track pedestrians in environments with only pedestrians present, such as parks or walkways (*Berclaz et al., 2006, Zhao & Nevatia, 2004, Isard & MacCormick, 2001*).

To handle environments with road users of different types is more complicated as it is also necessary to determine the type of each road user. Most systems have to be configured for each type of road user, typically by manually specifying a large set of length parameters of some wire-frame model or by training the system on a large amount of manually classified training examples (*Leibe et al., 2008, Schoenemann & Cremers, 2008, Berclaz et al., 2006, Messelodi et al., 2004, Zhao & Nevatia, 2004, Isard & MacCormick, 2001, Tan et al., 1998, Koller et al., 1993*). Other methods work with more coarse models where it is enough to specify some approximate size of the road users (*Song & Nevatia, 2007*).

The detection and tracking of cyclists and pedestrians are a bit different from vehicles in that they do not keep exactly the same shape when moving. The current state of the art for pedestrian tracking uses techniques such as bag-of-words, gradient-histogram or randomized-forests (*Leibe et al., 2008, Moosmann et al., 2008, Dalal & Triggs, 2005*). Less work has been done on detecting cyclists, and for many of the approaches the results are provided only for test images not produced from the typical surveillance angle, i.e., where the scene is viewed from above (*Moosmann et al., 2008, Agarwal & Triggs, 2006*).

Some efforts have been made to analyse videos recorded from vehicles (*Corneliu & Nedevschi, 2008*) or near-ground cameras (*Parkhurst, 2006, Xu et al., 2005*), aircrafts (*Reinartz et al., 2006, Zhao & Nevatia, 2003*), infra-red cameras (*Kirchhof & Stilla, 2006*) and complementing video with data from other sensors (*Wender & Dietmayer, 2007*).

The output of the video analysis algorithms in the form of trajectories, speed profiles, etc., provides description of individual road-user behaviour on a micro-level, which is a key input for many applications like assessment of safety and efficiency of traffic systems, evaluation of different road design solutions, calibration of behavioural models, etc. However, the reported attempts to use video analysis for purposes more sophisticated than just traffic counting and detection of simple incidents (congestions or movements in prohibited areas or directions) are quite limited. *Parkhurst, 2006*, makes an attempt to analyse the shapes of speed profiles, for example, to distinguish between road users who come to a complete stop or fail to stop in front of a stop-sign at an intersection. *Atev et al., 2005*, describe a method for predicting possible collisions by extrapolating the road users' trajectories, which can be used for early collision warning systems. *Messelodi & Modena, 2005*, propose an aggregated accident risk index which is calculated by utilising information about temporal and special relations between individual road users. Earlier, some validation of traffic-conflict techniques was done using semi-automated video systems (*Asmussen, 1984, Grayson, 1984*). *Saunier & Sayed, 2007*, extend the traditional approach to counting traffic conflicts and propose a method for estimating the probability of any

interaction between road users resulting in a conflict. *Zhang et al., 2007, Archer, 2005, Hoogendoorn et al., 2003*, are some examples of using video data for calibration of the micro-level behavioural models, while *Ervin et al., 2000*, create a database of vehicle trajectories that represent the natural driving environment and may be used, for example, to estimate the potential impact of new driver assistance technologies.

3.3. Video analysis system at Lund University

The conceptual scheme of the automated video analysis system developed at Lund University is shown in Figure 6. Three main blocks in system work may be distinguished – video recording, video processing and traffic data interpretation. The digital video is recorded by one or several cameras installed in the traffic environment. Then, it is processed by special algorithms that detect road users and extract their trajectories and other relevant parameters like speed, size and orientation, expressed in road-plane co-ordinates. This data is further interpreted, based on some traffic-related criteria (e.g. a certain type of manoeuvre or interaction, intensive braking, acceleration, etc.), and the relevant situations are detected. These situations create the raw material for the behavioural analysis itself and, in an ideal case, each situation is described by detailed data on trajectory and speed of the involved road users, illustrated by a video sequence from the original video and, possibly, complemented with other data like measurements taken with other types of equipment, on-site observations, etc.

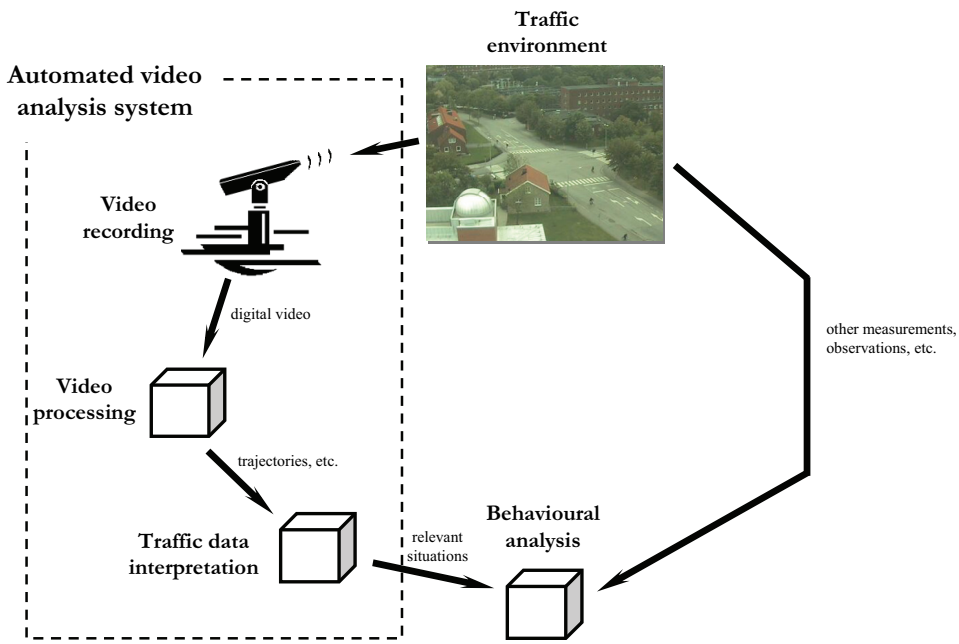


Figure 6. The conceptual scheme of the automated video analysis system and its place in a behavioural study.

The "borders" of the three system blocks are quite soft. For example, some initial processing of a video may be performed, while it is being recorded, by a processor embedded in the camera. On the other hand, many of the traffic interpretations may be carried out during the video processing phase, without a transfer to the road-plane co-ordinate system (e.g. detection of simultaneous presence, certain manoeuvres, etc.).

The video processing part of the system is built on the principle of a toolbox where a proper technique may be chosen and applied, depending on the study purpose and requirements as to the accuracy, processing time, etc. The techniques vary in degree of automation, complexity and computation intensity and, generally, the more advanced a technique, the more sensitive it is to eventual errors, quality of input data and calibration procedures and the more validation it requires. The input video is usually clipped in shorter files (between 15 and 30 minutes), which allows analysis of different files at the same time using several computers and thus decreasing the total waiting time for the results.

For the moment, the system includes the following algorithms:

- Advanced road user detection;
- Trajectory extraction I (interest points);
- Rectification;
- Speed estimation;
- Trajectory extraction II (Hidden Markov Model).

3.3.1. Advanced road user detection

This detector finds situations where some objects are moving within a certain area and in a certain direction, which allows the setting of more advanced detection criteria compared to virtual loop detectors. The detector can be used as an initial filter that removes a lot of uninteresting parts from the original video. For each of the detections a short video clip or a link, indicating the start and the end times in the original file, is saved so that it can be further analysed or manually sorted afterwards.

In the current version the detector is based on a KLT (Kanade-Lucas-Tomasi) interest point tracker (*Shi & Tomasi, 1994*). First, this algorithm finds points in the image that are expected to be easily found in the following frames, typically ones with a lot of structure such as corner points or edge junctions. These points are tracked over the entire video sequence, and as some points are lost new ones are found. Typically, several interest points are found in a single road user and large road users contain more interest points than small ones. Some results from this tracker are shown in Figure 7.

The tracks provide only approximate information about the road users in the scene. The track shape gives some idea of the road user's trajectory (in image co-ordinates) and the number of interest points moving simultaneously and close to each other can indicate the road user's size. By setting some heuristic rules, like location of the tracks, their direction, minimal and maximal number of interest points going together, etc., the situations of interest may be selected. Typically, the parameters are chosen with a

large margin so that no relevant situations are removed, even though it implies quite a lot of irrelevant situations among the detections.

a) b)

Figure 7. Results from the KLT interest point tracker** (Shi & Tomasi, 1994)**: a) a single frame with interest points found in it (shown in green); b) the tracks of interest points generated during a 30 min sequence** (*from *Laureshyn et al., 2009*)**.

3.3.2. Trajectory extraction I

This technique detects and follows road users in the video, providing detailed description of their trajectories in image co-ordinates.

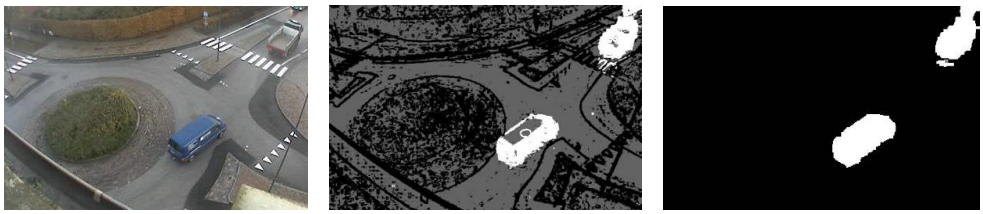
First, a foreground/background segmentation is done, i.e. it is decided which parts of the image currently show the road environment without moving objects in it (background), and which part shows the objects coming into scene and moving across it (foreground). A background model is created for each 8x8 block of pixels and contains a temporal mean and variance for each pixel, which are then compared to the corresponding pixels in each frame by using correlation coefficients. These coefficients are independent of intensity level, which makes the results fairly robust in case of changes in the lighting conditions (such as when a cloud shadow comes over a certain part of the image) as long as the change occurs over the entire 8x8 block. The output of this stage is a probability for each pixel to be foreground (between 0 and 1).

Then, the single pixels detected as foreground are clustered into larger components by using information about the surrounding pixels. By setting a minimal threshold for a component size, a lot of noise is removed at this stage. It also allows a single object to consist of several components, which occurs, for example, when it is partly occluded by a lamppost or something similar. Some results are shown in Figure 8.

The simplest way to estimate the position is to use the middle point of the component(s) that represents a road user in the image. However, this position estimate is very "jerky", as the pixels on the borders may be detected as foreground in one frame and as background in the next one, thus affecting the middle point position. The interest points used by the detector (described above) are more stable and may be followed over a longer sequence of frames. The connected components are used to cluster the interest point tracks into clusters of tracks belonging to the same road user. By calculating the mean over the interest points in each frame a point close to the centre of the road user is found, which is used as the resulting position of the generated track.

The extraction of trajectories is quite computationally intensive and much computation time can be saved by applying it only to the video sequences selected by the detector. However, the length of the processed video clip has to be at least 2-3

minutes long (to ensure the quality of the background model), and if the detected events are relatively frequent, the clips may "overlap" and in this case it is more reasonable to process the entire video film directly.



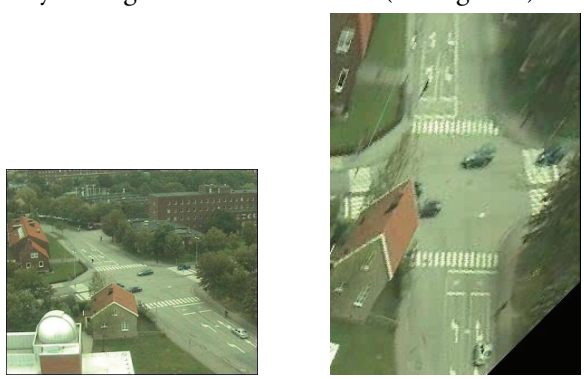
a) *b)* *c)*

Figure 8. Background/foreground segmentation and clustering: a) input frame; b) foreground probability per pixel (white pixels have high probability of being a foreground while the black pixels have a very low probability; grey colour indicates the probability close to 0.5, i.e. the model cannot decide whether it is foreground or background); c) resulting clustering (from *Laureshyn et al., 2009*)**.**

3.3.3. Rectification.

Transformation of the position of an object in the image to the real-world co-ordinates is not a simple task. The problem is that an image is a 2-dimensional projection of the 3-dimensional world and some information (image "depth") is missing. However, having some prior information about the scene enables the transformation to be done. For example, if the image is taken from directly above, it will resemble a map of the scene and the real-world distances may be calculated by simple scaling.

In reality, however, it is very seldom that a camera is located straight above the road but rather on top of a nearby building or on a mast standing at the roadside. It is still possible to find the transformation function between a certain plane (usually, the road plane) in the image and its representation in the real world (as explained in Figure 5). Modifying the image so that the image co-ordinates can be transferred to the real world co-ordinates by scaling is called rectification (see Figure 9).



a) *b)*

Figure 9. Rectification: a) the original image; b) the same image rectified. Note that the road plane is rectified nicely while objects above the road get more and more distorted the higher they appear (from *Laureshyn & Ardö, 2006*)**.**

The problem here is that it is not known which points belong to the road plane in the image and which do not. If a camera is high enough above the road plane, an assumption that all the road users are "flat" and lies on the road plane may be made. In this case the position of the road user, found as a mean of its interest points, may be assumed to be the middle point of the road user's "footprint" on the road and transferred to the real-world co-ordinates.

3.3.4. Speed estimation

The calculation of a road user's speed as a differentiated position, estimated as a mean of the interest points, yields quite inaccurate results since the position jumps a bit back and forward as new points are found and the old ones are lost. This does not happen if the set of interest points is considered to be dynamic, i.e., all the points might not be available in all the frames (*Åström et al., 2007*). A mean shape for the entire set of interest points is estimated and then its configuration in each frame is expressed by translation, rotation and scaling parameters that transform the mean shape into the shape observed in the image. Figure 10 shows an example of such transformation between two sets of interest points. The position of the road user is estimated by applying the same transformations to the centre point (mean of all the points) of the mean shape. This model does not take into account the fact that the interest points come from a 3-dimensional object projected into a 2-dimensional image, and therefore the estimated position might quite often appear outside the road user's borders. Still, differentiation of this position provides much smoother and accurate speed estimates.

Figure 10. Transformation between two sets of interesting points (from *Laureshyn et al., 2009)***.***

3.3.5. Trajectory extraction II

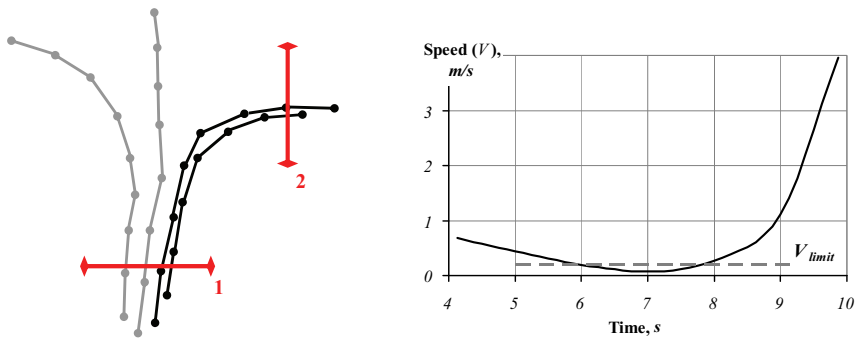
This technique uses another approach to estimating the positions of the road users based on a model where all the road users are approximated to a set of 3-dimensional "boxes" of pre-defined size and the entire analysed scene is described by its state consisting of the location, orientation and type of all the road users present at each time moment. A set of state hypotheses is generated and the probability of each is calculated by matching it to the observations made by the camera. The states are combined into sequences by designing a Hidden Markov Model (HMM), which is optimised over state sequences to find the one that describes the observations best. This state sequence contains the trajectories of all the road users (*Ardö, 2009*).

The technique has many advantages. It is possible to combine data from several cameras, which allows installing them at locations where the view is not so good (it is sufficient that a camera covers just a part of the studied area as long as the coverage

between the cameras overlaps). The technique is more robust in tracking and, if such a criterion is set, guarantees that no trajectories start or end in the middle of the scene. Since road users are represented by 3-dimensional boxes, there is no need to make the assumption that they are "flat" and thus introduce an error in position estimation. However, as the number of possible states (which depends on the number of considered road user types, cell size of the positions grid, etc.) is increasing, this technique is becoming more and more computationally intensive.

3.3.6. Interpretation of data extracted from video

In typical output from video processing algorithms, each road user is represented by a set of values describing position, speed and possibly other variables (acceleration, direction, etc.) with a certain time frequency (equal to frame rate) and during a certain time period (as long as the road user appears in the camera view). For many applications the analysis of such data is quite straightforward. For example, road users moving in a certain direction may be selected by setting some "gates" that a trajectory has to pass through in the right order (Figure 11*a*). Simultaneous arrival and possible interaction may be detected by comparing the time at which road users enter and leave the scene, possibly complemented with the analysis of directions of movement as well. The detection of a stop usually requires setting a minimal speed limit by which a road user is classified as "stopped" since, due to inaccuracy in the speed measurements, the speed is almost never equal to zero but fluctuates around it[1] (Figure 11*b*). Many simple indicators like lateral position, distance to a stop line, average travel speed, etc. may also be calculated easily.



a) *b)*

Figure 11. Some examples of simple detection techniques: a) right turning vehicles are selected if their trajectories pass through gate 1 and then gate 2; b) a typical speed profile for a vehicle that stops for a short time to yield at an intersection and then continues moving. Note that the measured speed never comes down to zero, but if the speed threshold (V_{limit}) is set slightly above zero the stop will be detected.

The data produced by the video analysis algorithms has the potential to be used for more advanced analysis. The great advantage is that it provides a description of continuous processes taking place in traffic. Studying the process rather than its state

[1] This problem is discussed in *Andersson, 2000.*

at a certain moment allows for better understanding of how road users interact and what factors affect the interaction at different stages. Sequential data allows the calculation of more complex indicators. For example, from a speed profile it is possible to calculate acceleration and the second derivative of speed, i.e., jerk. This cannot be done if speed is measured just at a certain point. Another problem with using point measurements is that it is necessary to know for which moment or location the measurements should be taken so that it will be representative for the entire interaction. This decision requires a lot of validation work and it is a great advantage to have a continuous description so that values at different moments can be tested and compared.

4. STUDYING THE TRAFFIC PROCESS

This chapter further explores how the advantages of the sequential data extracted from video can be utilised. Section 4.1 considers studying the interactions between individual road users from a safety perspective and describing the safety situation at a certain site using operational severity hierarchies. Section 4.2 deals with the application of pattern recognition techniques for classifying the sequential data, an important step in distinguishing between different behaviour types and defining of the "normal" and "critical" traffic processes.

4.1. Continuous safety indicators

4.1.1. Severity rating with time-based indicators

In the discussion on the indirect traffic safety indicators in Chapter 2.2, I argue that if one can organise all the elementary events in traffic into a severity hierarchy, the shape of the hierarchy can be used for description of the safety situation and also the relations between safety and other qualities like efficiency, comfort, etc. The severity of an event is defined as a measure reflecting its probability of developing into an accident and the severity of consequences *if* this occur. The problem, however, is to find indicators that reflect this measure in an objective way.

Several approaches to the severity estimation have been discussed in the literature (*Nygård, 1999, van der Horst, 1990, Hydén, 1987, Grayson, 1984, Allen et al., 1978*). These include proximity (in time, space) between road users, intensity of a necessary evasive action (e.g. braking), road users' speeds, etc. Most of these indicators reflect just "one side of the truth". For example, an encounter at a higher approaching speed is probably more severe than at a low speed, but even at a high speed an encounter can be successfully resolved if road users have enough time and space. Similarly, a close distance between road users does not say much if one does not know what their speeds are.

In this respect time proximity is a bit special as it reflects, in some way, both the speed and the distance proximity. Probably, for this reason, many of the existing traffic-conflict techniques use some kind of time-proximity indicators as a basis for conflict detection, in some cases complemented with other measures as well (*Hydén, 1987, van der Horst & Kraay, 1986, Grayson, 1984*).

One of the first studies where time proximity was used for rating traffic conflicts with regard to their severity was carried out in the USA in the 1970s. *Hayward, 1971*, (cited *Hydén, 1987*) proposed an indicator called Time-Measured-to-Collision (TMTC) defined as "the time required for two vehicles to collide if they continue at their present speeds and on the same path". Later, the name was changed to Time-to-Collision (TTC). TTC has some important properties. First, it exists only if road users are on a collision course (i.e., if they continue without changes, they will collide). It cannot be measured directly but is rather calculated by predicting the road users' motion. TTC is a continuous indicator and may be calculated for any moment as long as the road users are on a collision course. The theoretical TTC curve discussed by Hayward is shown in Figure 12. The curve starts when the road user gets onto a collision course, reaches some minimum value (zero in case of a collision) and then "jumps" into infinity when the collision course stops existing.

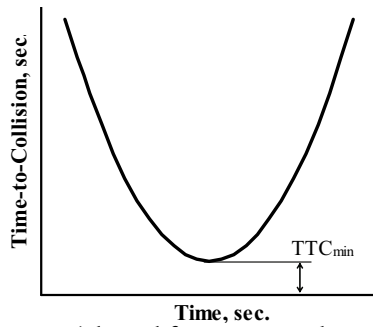


Figure 12. Theoretical TTC curve (adopted from *Hayward, 1971*, cited *Hydén, 1987)***.**

My minor comment is that the TTC curve does not necessarily have to "jump" at the end. This is true when the collision is avoided through braking by one of the road users. It is not the case in such situations as shown in Figure 13. The TTC is defined by the distance S and the speed of the vehicle 2, v_2 ($TTC = S/v_2$), i.e., the speed of vehicle 1 does not affect the TTC value as long as the vehicles are on a collision course. However, if vehicle 1 accelerates above a certain limit, the collision course stops to exist, i.e. vehicle 1 will leave the conflict area before vehicle 2 arrives at it. The TTC curve will still go down as long as the vehicles are on a collision course, but the moment the speed of vehicle 2 becomes high enough to avoid the collision, the TTC simply ceases to exist (note that it is not the moment vehicle 1 leaves the conflict area; the collision course ceases to exist earlier).

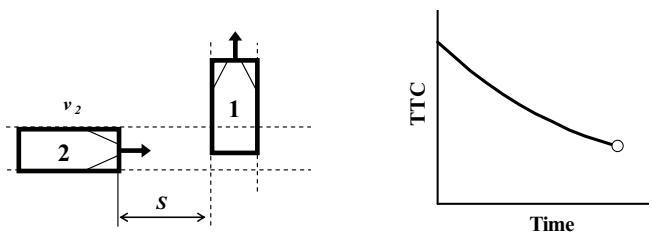


Figure 13. An example when TTC does not become infinity when the collision course ceases to exist.

Hayward collected the entire TTC curve, but used only the minimum TTC value (TTC_{min}) in analysis. TTC_{min} is an important point as it characterises the moment the road users come closest in time to each other. However, this is not the only possible indicator that can describe the TTC curve. *Hydén, 1987*, proposes using the TTC value at the moment the first evasive action is taken by one of the road users (this value is called Time-to-Accident, TA). Hydén puts forward several arguments for the use of TA. First, it is more operational as it is easier for an observer to define the moment for which the TTC is to be calculated, while to find TTC_{min} the entire TTC curve is necessary. Note that, on the other hand, finding the moment of the start of an evasive action by looking only at the TTC curve or, for example, the speed profiles of the road users is not always easy, especially if the evasive manoeuvre is other than braking. TA represents a critical moment in the whole course of events when an unexpected hazard is detected. From a theoretical point of view TA is more universal as it can be applied to events that result in a collision (accidents) and to events where collision is avoided. TTC_{min} becomes equal to zero in the case of a collision, thus not allowing us to distinguish the severity of the events that finally become accidents.

An encounter is a continuous process while both TTC_{min} and TA are measures of certain moments in it. The question is what moment in the entire process characterises the severity in the best way, and whether the best characteristic is actually a momentarily value of TTC or some characterisation of the process, for example the lengths of the time spent on a collision course or the slope of the TTC curve. It is easy to show that when the momentarily values are used, quite different situations can be classified as if they are very similar. One example is given in Figure 14 (adopted from *Hydén, 1987*). A pedestrian is crossing the road in front of a car. Several different scenarios are possible for this situation to develop starting from the moment marked as point A on the TTC graph (the distances and the speeds are also given for this moment):

a) Both keep the same speed and path and collide in 1.1 seconds (the TTC curve follows line I, the moment of the collision is point B);

b) The pedestrian accelerates slightly and leaves the conflict area before the vehicle arrives. TTC reaches the minimum and ceases to exist at point C;

c) The driver applies the brakes for a short time, the TTC curve follows line II and the pedestrian, again, leaves the conflict zone before the vehicle (TTC curve ends at point D);

d) The pedestrian "freezes" in front of the car and the driver brakes very hard and stops in front of the pedestrian. In this case the TTC will follow line III, reaching the minimum value at point E.

In the last three scenarios the evasive action starts at point A, i.e., TA is the same in all the cases and equals 1.1 seconds. The minimum values on the possible TTC curves (points C, D, and E) are also quite the same (around 0.2 second). Thus neither TA nor $TTC_{min,}$ or even both of them used together, enable us to discern that the situations are quite different.

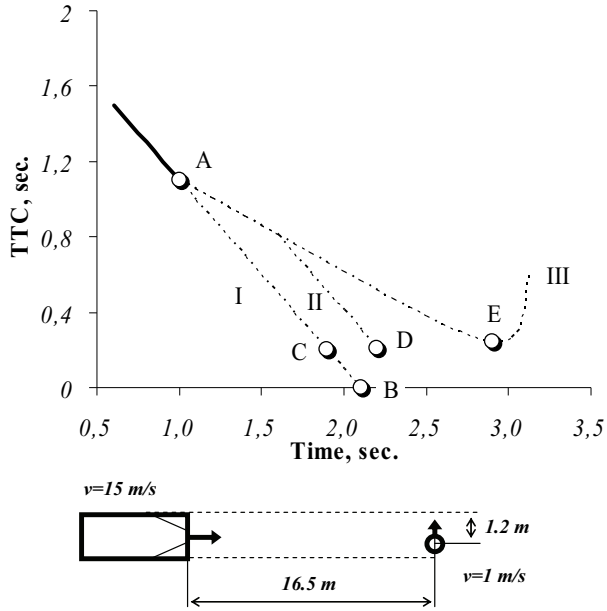


Figure 14. Different scenarios for development of an encounter between a pedestrian and a vehicle described with TA and TTC_{min} (adopted from Hydén, 1987).

To overcome this problem, *Minderhoud & Bovy, 2001*, propose two enhanced TTC-based indicators. The first one, Time Exposed TTC (TET) is the time during an encounter when the TTC is below a certain threshold value, TTC^* (Figure 15). TET reflects the duration of the most critical part of an encounter when TTC is low.

The second indicator is called Time Integrated TTC (TIT) and calculated in the following way:

$$TIT = \int \left[TTC^* - TTC(t) \right] dt \,, \qquad 0 \le TTC(t) \le TTC^* \,.$$

TIT is the area between the threshold level TTC^* and the TTC curve when it goes below the threshold, thus reflecting both the lengths of the time with low TTC and the extent to which the TTC sinks below the threshold.

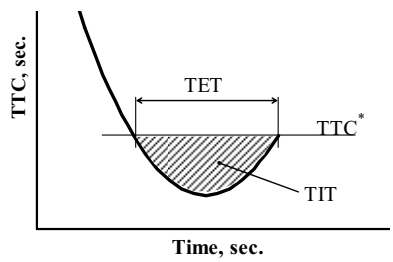


Figure 15. The definition of TET and TIT (adopted from Minderhoud & Bovy, 2001).

The problem with these indicators is that they can only be applied to critical situations with low TTC, i.e. the "normal" encounters are excluded. Even if they are to be used only for analysis of conflicts, the next problem is selection of a proper

threshold level. Different studies propose different thresholds to distinguish between critical and "normal" events, for example, Minderhoud & Bovy use 3 seconds in car-following interactions on a motorway (*Minderhoud & Bovy, 2001*), Hayward – 1 second for car-car conflicts (*Hayward, 1971*, cited *Hydén, 1987*), Hydén and van der Horst – 1.5 second for any types of traffic conflicts (*van der Horst, 1990, Hydén, 1987*). In the latest version of the Swedish Traffic Conflicts Technique Hydén also makes the threshold between serious and non-serious conflict dependent on the speed (*Hydén, 1987*).

Collision course is an absolutely necessary condition for a collision to occur. However, the situations without a collision course but close to it are also of interest since even a small adjustment of road users' course or speed may put them on a collision course and make collision possible. Observations of road users in such situations suggest that they still behave *as if* they are on a collision course (*Svensson, 1998*). *Hydén, 1987*, hypothesises about some kind of safety margin used by road users when judging if they are on a collision course or not. There is still too little knowledge available about exactly how time proximity is perceived by road users and if it is the time proximity that is used for judging the severity of the situation. Some studies indicate that road users are able to retrieve TTC information directly from the optic flow field, and that this information is important in decisions about being on a collision course or not, the start of braking and control of the braking process itself (*van der Horst, 1990*).

The main drawback of TTC is that it cannot be calculated without a collision course, even if the road users miss each other by a very small margin. *Svensson, 1998*, solves this problem practically by slightly increasing the actual size of road users so that calculation of TTC becomes possible. *Allen et al., 1978*, come up with another indicator for treating the situations of near-misses called Post-Encroachment Time (PET). PET is defined as the time between the first road user leaving the conflict zone and the second one arriving at it (Figure 16).

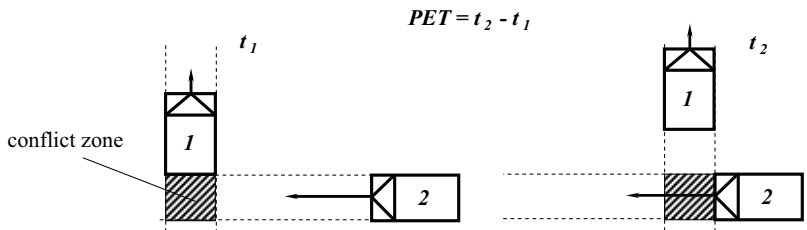


Figure 16. Conventional definition of Post-Encroachment Time (PET).

There is just one value of PET for an encounter and it can be measured directly with a stopwatch. A practical advantage of PET is that to measure it one needs to observe only a small area around the conflict zone (for TTC the road users have to be observed several seconds before they arrive at the conflict zone). On the other hand, being originally developed for studying left turns on rural roads, PET does not perform very well in urban conditions as it cannot be applied in cases when one of the road users stops or when the road users are on a following course (*van der Horst, 1990*).

Since PET is just a single value describing a continuous process, the argumentation about limitations of TTC_{min} and TA applies to PET, too (i.e. encounters very different in character can end up with the same PET value, which does not allow us to distinguish between them in any way). The concept of PET can be broadened to a continuous parameter which says what the PET value is expected to be for each moment if the road users continue with the same speeds and paths. To my knowledge, this indicator has been used very little in traffic safety research, but it is quite common in studies on traffic flow theories and gap acceptance distributions (for example, in *Hansson, 1975*, where it is called Time Advantage; I will use the abbreviation TAdv).

While TTC reflects the closeness to a collision point when on a collision course, TAdv says something about closeness to a collision course. The specific of Time Advantage is that while its low values may reflect the safety aspects, the higher values (above 2-3 seconds) describe the normal traffic conditions and may be seen as a measure of one road user's power (advantage) over the other in a competition over the same spatial zone. A road user having a large time advantage is most likely to be the one to pass the common zone first. However, if the time advantage is small, the second road user may accelerate with the aim of passing first instead, which occurs primarily when one of the road users is "stronger" than the other, for example, in the case of a private car vs. a pedestrian (*Várhelyi, 1998*) or a truck vs. a private car. The important point here is that the use of the same indicator to describe both safety and efficiency of the traffic processes has certain advantages and may help to better understand how these two qualities are balanced by the road users and to verify the hypotheses of such a relation (*Svensson & Hydén, 2006, Svensson, 1998*).

For the moment, TTC and PET are two separate concepts that have their "followers" and discussions about one indicator being better than the other are not unusual. However, using just one of the indicators means excluding the aspects reflected by the other one, which seems quite unwise. Real traffic includes situations both with and without a collision course, and even during the same encounter the transfer between the two types occurs smoothly. Generally, the relation between two road users at a certain moment can be classified as belonging to one of the three types (Figure 17):

Type A (collision course). Two road users are on a collision course and they will collide if no evasive action is taken;

Type B (crossing course). The road users' paths overlap, but collision will be avoided since they pass the common conflict zone with a time margin;

Type C (diverging course). Two road users are on parallel or diverging course and their paths do not overlap. This does not mean that the risk of a collision is completely zero since a change of the paths (often very minor) may create a common conflict zone and even put them on a collision course. For example, in a meeting on a two-way road even a small shift towards the middle of the road by one or both vehicles can result in a collision with quite severe consequences.

By adjusting their course or speed, or both, two road users can change the type of an encounter. However, their behaviour does not change abruptly at the moment of transfer and, most probably, neither does the severity. Therefore, the indicators used

to describe the encounter should also allow smooth transfer between the encounter types.

I propose using a set of indicators that allows continuous description of an encounter process and can handle the transfer between A(collision course) and B(crossing course) situations. The C-type (diverging course) situations are omitted for now. The reason for this is difficulties finding a good indicator that can be used. It may be argued, though, that for a collision to become possible at all the situation has to change first into B- or A-type, and when it happens the indicators developed for these types can be used. This does not mean that C-situations are not relevant; in the further development of the concept these situations have to be included.

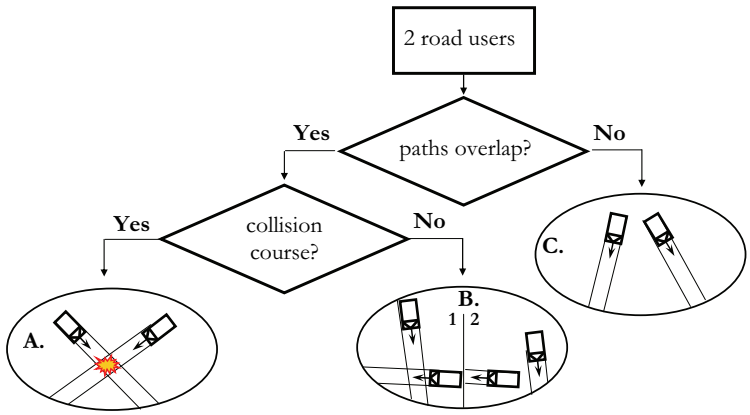


Figure 17. Classification of encounters based on spatial and temporal relations between the road users.

The proposed approach has to be seen as a first approximation that needs to be validated and further elaborated. So far there has been very few validation studies of the indicators describing interaction between road users and their relation to safety. My belief is that with automated video analysis more of such studies will be possible.

4.1.2. Description of an encounter process with a set of indicators

Time-to-Collision – a measure of the closeness to a collision when on a collision course

In most of the studies TTC is calculated using the simple assumption that the road users' trajectories cross at a right angle or are parallel (Figure 18). For example, *van der Horst, 1990*, calculates TTC for the case of right-angle approaching using the following equations:

$$\begin{cases} TTC = d_2/v_2, & if \quad d_1/v_1 < d_2/v_2 < (d_1 + l_1 + w_2)/v_1 \\ TTC = d_1/v_1, & if \quad d_2/v_2 < d_1/v_1 < (d_2 + l_2 + w_1)/v_2 \end{cases},$$

where d_1, d_2 are distances from the fronts of vehicles 1 and 2, respectively, to the area of intersection (Figure 18*a*);

l_1, l_2, w_1, w_2 are lengths and widths of vehicles 1 and 2, respectively;
v_1, v_2 are vehicles' speed.

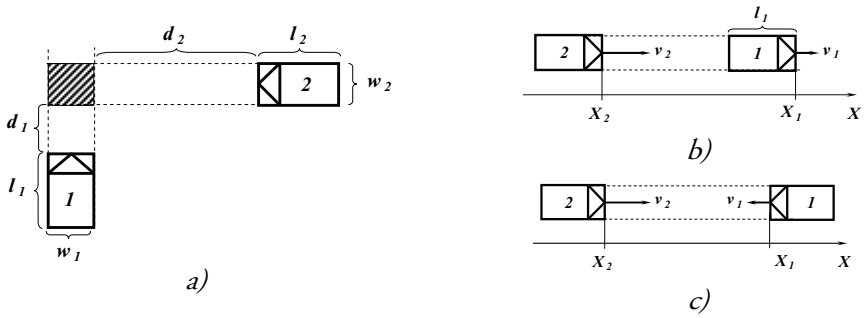


Figure 18. Calculation of TTC for perpendicular (a) and parallel courses (b, c).

For the case of rear-end collision (Figure 18*b*), *Minderhoud & Bovy, 2001*, calculate TTC as:

$$TTC = \frac{X_1 - X_2 - l_1}{v_1 - v_2}, \quad if \;\; v_2 > v_1,$$

where X_1 and X_2 are the positions of the front parts of vehicles 1 and 2, respectively.

For the case of a head-on collision (Figure 18*c*), the previous equation can be easily modified to:

$$TTC = \frac{X_1 - X_2}{v_1 + v_2}.$$

However, in a general case two vehicles can approach each other at any angle and, moreover, for the same angle different collision types are possible (Figure 19).

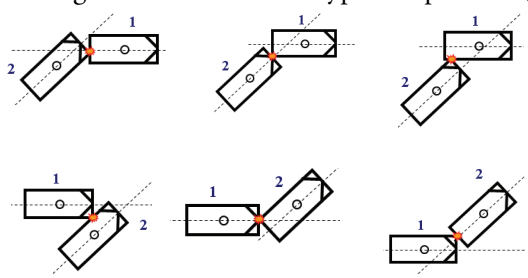


Figure 19. Possible collision types for the same approaching angle (adopted from van der Horst, 1990)

From Figure 19 one can conclude that it is always a corner of one vehicle that meets a side of the other one. Since in the general case it is not known which corner will meet which side, all possible pair-wise combinations have to be considered (i.e. 32 combinations assuming that road users have rectangular forms). If a collision is possible in several combinations, the lowest TTC-value among all the side-corner combinations should be taken, since it is this side and corner that will come into

contact first in case of a collision. The detailed calculation procedure can be found in Paper II.

Time Advantage – a measure of the closeness to a collision course

The conventional geometry-based definition of Time Advantage is also difficult to apply when vehicle trajectories do not cross at a right angle. The entrance and exit from the "conflict zone" are no longer time moments but periods, and it is even possible that both road users appear in the "conflict zone" at the same time but still avoid the collision (Figure 20).

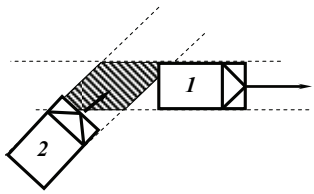


Figure 20. The problem with the "geometrical" definition of TAdv – both road users appear in the common zone but avoid collision.

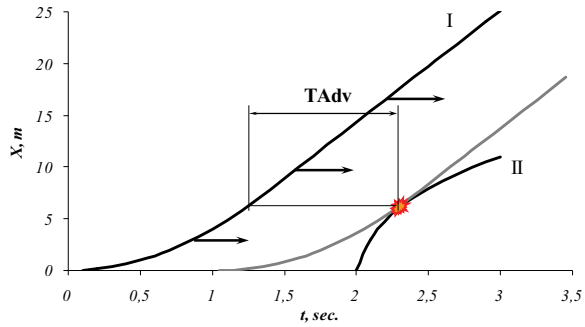


Figure 21. "Delay"-based definition of TAdv.

To overcome this problem other non-geometrical terms may be used. Basically, TAdv reflects the correction in time that is necessary to put the road users on a collision course. Therefore, I propose the following definition for TAdv: "the minimal delay of the first road user which, if it is applied, will result in a collision (assuming that otherwise the road users keep the same speeds and paths)". Figure 21 helps to explain this definition. Lines I and II describe the *predicted* movements of two road users over the time (for simplicity I consider only one dimension and neglect the physical size of the road users). The "delay" of road user I means that its travel line has to be shifted along the time axis until it touches line II. The length of the time shift here is the Time Advantage.

In practical calculations, when the dimensions of the road users are taken into account, the TAdv has to be calculated for each possible side-corner combination. For the same reasons as in the case of TTC, the lowest TAdv-value found should be used. The calculation procedure is provided in Paper II.

T_2 - a measure of the closeness to the conflict zone

Time Advantage itself is not sufficient to describe the probability of a collision since it is also important to know how soon the encroachment is to occur. Even if TAdv is small at a certain moment, the road users might have plenty of time to adjust their speeds and trajectories and increase it. Therefore some measure of how soon the road

users will actually arrive at the conflict zone is necessary. For this purpose I propose using the time of the second road user arriving at the expected collision point (position of which is calculated on the assumption that TAdv is applied to the first road user). This parameter is further abbreviated as T_2.

To use the second road user appears to be more safety-relevant since, whatever the actions of the first road user are, it is the one who arrives latest who has most time to take an evasive action. However, if the moment of the first road user leaving is of interest, it can be easily calculated as ($T_2 - TAdv$).

Interpretation of T_2 is very similar to TTC, i.e. it describes how soon the collision will occur if the road users come on a collision course. This similarity provides "smooth" transfer between the "collision course" and "no collision course" situations. At the moment of transfer from "collision course" to "no collision course" the TTC ceases to exist, and Time Advantage starts to increase from zero. At this moment T_2 is equal to TTC, and if both TTC and T_2 are plotted on the same graph, they will make a continuous curve. Similarly to TTC, T_2 will "jump" into infinity if the second road user comes to a complete stop.

Time Gap - a measure of spatial proximity

Depending on the relation between road users' trajectories and speeds, the collision point, for which TTC or TAdv are calculated, can be far ahead while the actual distance between the road users might be not as large. This is especially noticeable when the road users' trajectories are parallel or close to it (Figure 22).

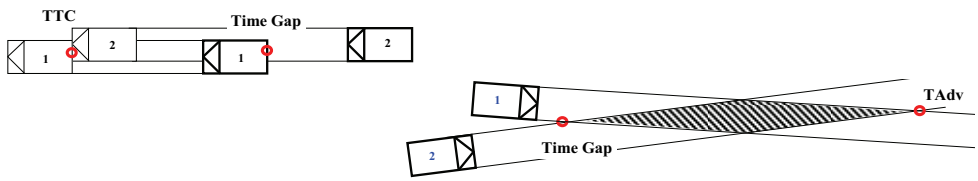


Figure 22. The points for which Time Advantage and Time Gap are calculated.

To reflect the actual spatial proximity (in time units), I propose to use the Time Gap (TG) indicator. Time Gap in its conventional definition is applied to vehicles following in a flow and is measured as the time between the moment the rear end of the first vehicle passes a certain point on a road and the front of the following vehicle arrives at it (*Vogel, 2002*). This definition is difficult to use if road users are not on a parallel course. To extend this parameter, preserving its main concept, the following definition is proposed.

Imagine that the first road user is delayed so that the road users get onto a collision course. Depending on the delay size, many collision points are possible. The delay that creates a collision point as close in time as possible is chosen. The Time Gap here is the time necessary for the second road user to arrive at this collision point. This definition includes the case of following on a parallel course (and in this case produces the same value as the conventional definition), but can also be applied to any cases of overlapping courses. In case of a following course, if the speed of the first road user is higher than the speed of the second one, TG becomes close to TAdv. When the road users approach at the right angle, TG gets close to T_2.

As it is not known in a general case which road user is "the first" and what type of collision is the nearest in time, all possible combinations of road users' sides and corners have to be considered. Again, the lowest TG value is to be used. The calculation procedure for TG is provided in Paper II.

Presumably, Time Gap has a weaker connection to collision risk compared to TTC, since TG considers the spatial proximity (in time units) only. It can be used, however, for detection of potential risks at earlier stages of an encounter, especially *before* the TTC can be calculated (i.e. without a collision course). This may be explained by an example of two vehicles moving on parallel courses at the same speed (no collision course, TTC $\rightarrow \infty$). If the first one starts braking, the vehicles will suddenly find themselves on a collision course and the pace of TTC decrease will depend highly on the size of the time interval between the vehicles (i.e. TG). Thus, TG reflects the probability of TTC quickly reaching low values *if* the road users get onto a collision course.

Speed

Even though time-based indicators combine and thus reflect both the spatial proximity and speed of the road users, some important information is still missing. This may be seen in a simple example. Consider two encounters with a collision course at the moment when TTC = 1.5 seconds, and in one of them the road users approach each other at speeds of 10 m/s and in the other at 20 m/s. To avoid a collision, in the first case the drivers have to decelerate by at least 3.3 m/s^2, while in the second case the deceleration has to be at least 6.6 m/s^2. Obviously, the second situation is more severe as it requires more intensive evasive action (braking) to prevent the collision.

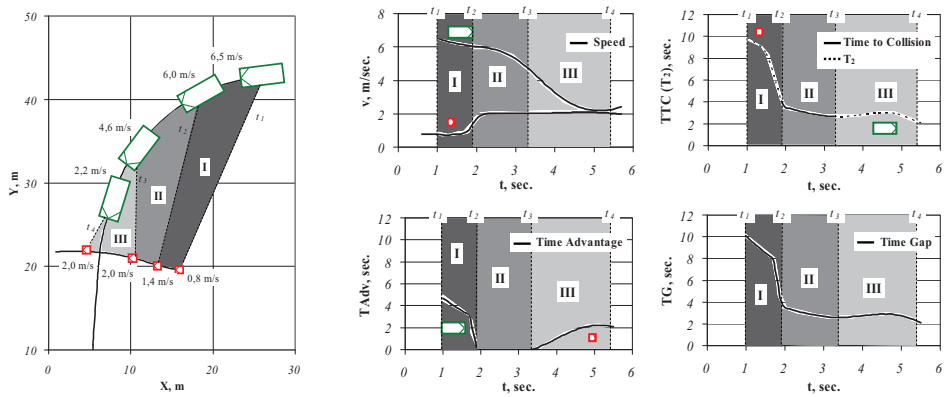
This example illustrates that the time-based indicators need to be complemented with some speed-related measure. The speeds of road users are very important since they affect both the probability of a collision *and* the severity of the consequences in case of a collision (as speed is strongly related to the collision speed and the amount of kinetic energy released). The Swedish Traffic Conflicts Technique (*Hydén, 1987*) uses Conflicting Speed (CS) as one of the parameters to define the conflict severity. The CS is defined as the speed of the *relevant road user*, i.e., the road user who takes the first evasive action, at the moment the action starts. This definition means that the speed of the second road user has no effect on the severity of the conflict, which is a limitation that has been much criticised. For example, in case of a vehicle-pedestrian conflict, the speed of the vehicle clearly has great importance for the conflict severity. However, as long as the evasive action is taken by a pedestrian, the majority of the conflicts are classified as having low severity due to the low speed of the pedestrian.

Shbeeb, 2000, specially studies the application of the Swedish Traffic Conflicts Technique in situations involving pedestrians and proposes using the higher of the two road users' speeds as CS, no matter which of them takes the evasive action. This definition makes the conflict score generally more severe, thus increasing the number of conflicts that are classified as serious can be and used in analysis. The correlation between the number of serious conflicts, according to the modified definition, and the number of reported fatal and injury accidents also improves.

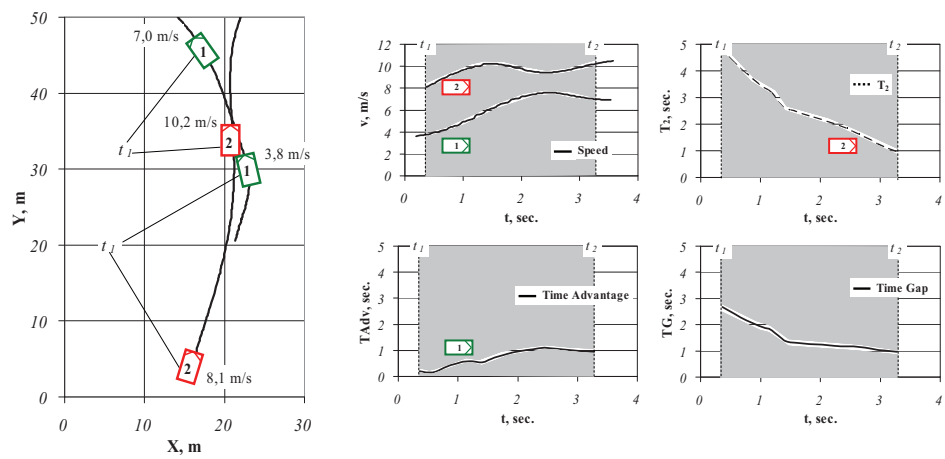
Still, this does not solve the problem with the speed of one of the road users being ignored. A possible solution is to use the speeds of both road users, taking into account the directions of the road users' motion (for example, calculating relative speed). Moreover, the process of speed adjustment during an encounter provides important behavioural information that can be used for classifying the situations and, possibly, distinguishing between normal and critical encounters (this is discussed in more detail in section 4.2). At this stage, I simply include the speed of both road users in the indicator set.

Two examples - crossing and following courses

Figure 23 illustrates how the proposed indicators are used to describe the interaction between two road users.



a) car-pedestrian



b) car-car (parallel course)

Figure 23. Two examples of parameter calculations for an interaction.

The first example (Figure 23*a*) describes an encounter between a car and a pedestrian at a pedestrian crossing. First, the car has a time advantage (i.e. is about to pass first)

as the pedestrian hesitates and keeps a very low speed (phase I). Then, however, the pedestrian, who has priority according to the traffic rules, decides to go first and increases speed to a normal pace. The TAdv of the car goes rapidly down to zero and from moment t_2 they appear on a collision course (phase II). TTC is decreasing as they approach each other. Noticing the pedestrian's behaviour, the driver brakes and from moment t_3 they are no longer on a collision course and TAdv (now the pedestrian's) starts gradually growing from zero (phase III). From moment t_4 the pedestrian is no longer in the way of the car and none of the indicators can be calculated. In this example the TG curve is quite close to the TTC and T_2 curves and does not contribute much additional information.

In the second example (Figure 23*b*) two cars move one after the other in a ring of a roundabout. The speed of the following car (marked as 2) is higher and if it continues to follow car 1 they will collide. However, car 2 drives off the roundabout, while car 1 continues in the ring. The expected collision point (for which the TAdv is calculated) lies in the area of the trajectories' divergence. Here the distance between the two cars is shorter than the distance to the collision point and the TG curve goes lower than T_2.

4.2. Behaviour classification by pattern recognition techniques

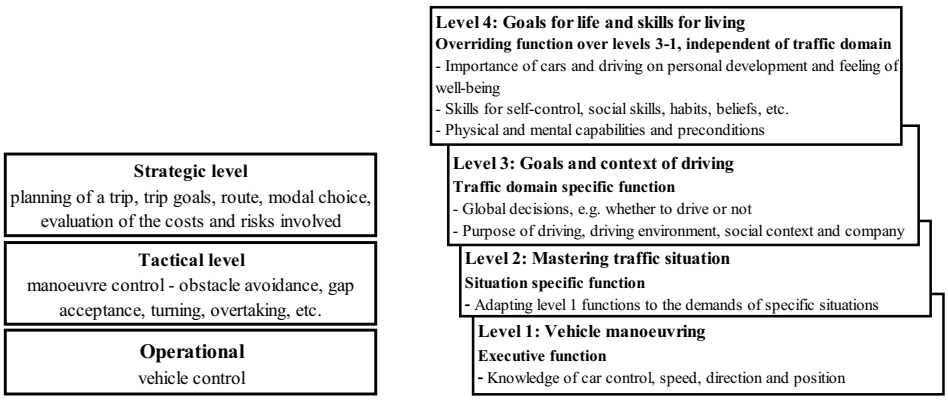
4.2.1. From operational to tactical data

Psychological theories that explain the performance of a road user in traffic often refer to a certain hierarchy of the tasks and decisions that have to be taken at different levels. For example, *Michon, 1985*, describes the following model. The first level (see Figure 24*a*) is operational and relates to control of the vehicle and decisions about the use of the steering wheel, pedals, gear choice, etc. The second, tactical level refers to manoeuvring and immediate interactions with other road users. The third, upper level, in the hierarchy is strategic and concerns tasks like trip planning, navigation and route choice. A more recent model, the GADGET-matrix (named after the European project for which it was developed, see Figure 24*b*), suggests a fourth level, described as "goals for life and skills for living" and referring to social skills, beliefs, importance of driving for personal well-being and social status, etc. (*Peräaho et al., 2003*).

From this perspective, the road user's momentary position, speed, etc. belong to the lowest operational level. The behaviour, i.e. road user actions in order to adapt the position and speed to the current traffic situation, belongs to the second, tactical level. In other words, the characteristics of behaviour are to be found in the changes of trajectory and speed profile shapes and their relation to the factors in the traffic environment. The challenge, however, is to extract the important behavioural information from the extensive data collected at the operational level.

Analysis of speed profiles of the individual road users is one of the most typical examples of such studies, and several approaches can be found. One way is to use a qualitative description of the road user's motion. An observer makes a note, for

example, that a road user "slows down", "yields", "drives on red", etc. (e.g. *Hakkert et al., 2002, Carsten et al., 1998*). The use of human observers puts serious limitations on the amount of data that can be practically collected and also on how detailed the data could be. On the other hand, humans judge the observed situations holistically, and in classifying them might consider dimensions not even captured by objective measurements.



a) *b)*

Figure 24. The hierarchy of the road user's tasks: a) adapted from** Michon, 1985**; b) Gadget-matrix** (Peräaho et al., 2003)**.

Another tradition is to interpret the speed data, collected by speed loggers installed in vehicles, in terms of driving patterns. This approach is used, for example, to produce standard driving cycles for vehicle tests and estimation of emission factors (*Larsson, 2009, Ericsson, 2000, André et al., 1995*). The problem, however, is that the studied population of drivers and vehicles is normally very limited and it is often difficult to relate the data to the actual traffic situation (e.g. presence of other road users).

Video analysis provides position and speed data with high time frequency and for large populations of road users. However, the problem of interpretation in behaviour-terms still remains, and the greatly increased amount of data that is collected necessitates automation of the analysis.

Simple aggregation to an average speed profile (*Várhelyi et al., 2004, Karlgren, 2001, Hydén & Várhelyi, 2000, Várhelyi, 1998*) loses information on differences between the individuals and variety of the individual shapes. It may therefore result in aggregation bias and a misleading final shape if the individual speed profiles are very different in character. Proposed by *Sekine & Sekine, 2009*, LUNA (Location UNiversal Archive), an aggregation format provides speed distributions at several points along the studied section. This preserves to a greater extent the variation of speeds at each cross-section point, but the longitudinal connections between the points of the individual profiles are lost. Neither of these approaches utilises the information on the variation in the shapes of individual profiles, which can be attributed to different behavioural strategies.

In order to utilise the advantages of the detailed data contained in large samples of speed profiles, it is necessary to have a method that:

- differentiates between the behaviour types based on endogenous (derived from data) criteria in a way similar to a human observer;
- makes use of the systematic variation in the data that can be attributed to different types of behaviour (i.e. analyses shapes of the speed profiles);
- can handle large amounts of data as produced by the video analysis techniques.

In this section I describe how this methodological gap can be filled by using pattern recognition techniques. Pattern recognition is a topic in machine learning theory that aims at classifying data based on *a priori* knowledge or information extracted from the data itself. Three techniques – cluster analysis, supervised learning and dimension reduction – have been tested.

The more extensive description of these tests can be found in Paper III.

4.2.2. Pattern recognition at work

The dataset: left turning vehicles at a signalised intersection

The data for this example was collected at a signalised intersection in Lund (Figure 25). The vehicles making left turns from one of the entrances were detected and their speed profiles saved. The data was visually checked for consistency and manually corrected in cases of obvious errors in detection. The profiles were trimmed and adjusted so that each profile contained the same number of points, evenly distributed along the trajectory between the defined start and end lines.

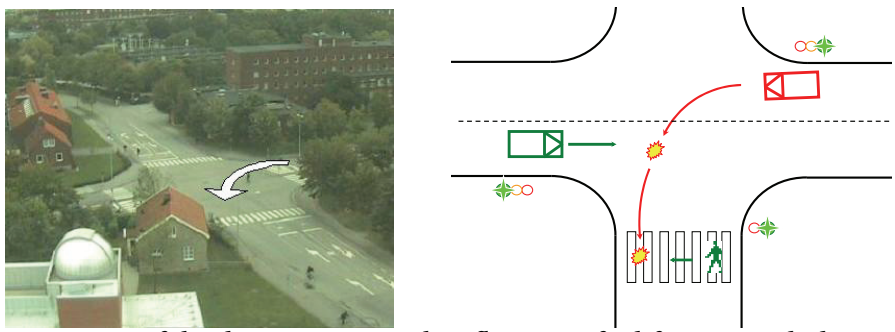


Figure 25. View of the observation site and conflict points for left turning vehicles.

According to the rules, a left-turning vehicle must yield to traffic coming from the opposite direction and to pedestrians that have green in the same phase. As a result, four situation types are possible. The first one takes place when there are vehicles coming from the opposite direction and the driver has to yield by braking in the middle of the intersection (near the imaginary middle line – type *a*). If there is a pedestrian at the pedestrian crossing, the driver has to brake before the crossing (type *b*). If no conflicting traffic is present, and the speed of a turning vehicle remains nearly constant or slightly increases (type *c*). Situations when a driver has to brake both near the middle line and near the pedestrian crossing are extremely rare, since

the pedestrian flow is low and those who are present usually manage to complete their passage while the driver is waiting at the middle line.

Examination of the vehicle speed profiles in such situations reveals that they have quite typical shapes in general (Figure 26). However, not all the profiles resemble the typical shapes, appearing somewhat in between two shapes, which makes it difficult to assign them to a certain type. A similar problem is experienced by an observer who classifies the situations by watching them on video, as they seem to fit the definition of more than one type (for example, a vehicle moves forward slowly to the middle lane, thus avoiding abrupt braking but is still affected by the oncoming traffic). This may be seen as a natural variety of the behavioural forms, which complicates classification regardless of what method is used.

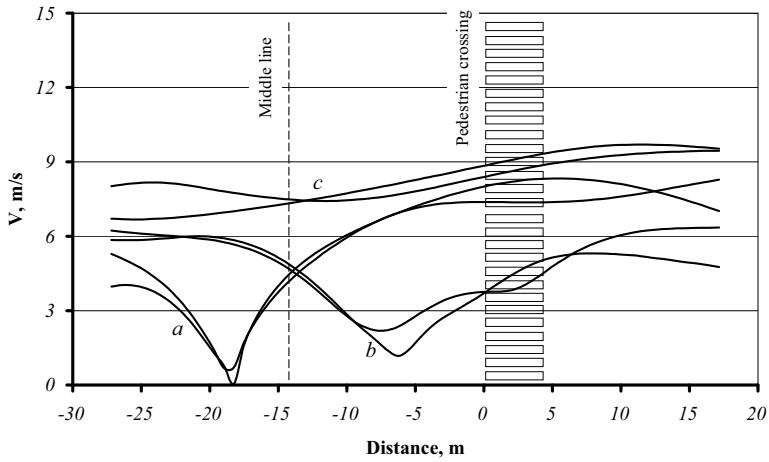


Figure 26. Three typical profile shapes: a – driver yields to the oncoming vehicles; b – driver yields to pedestrians at the pedestrian crossing; c – no on-coming traffic or pedestrians.

The techniques: cluster analysis, supervised learning and dimension reduction

To classify the speed profiles, I use three basic pattern analysis techniques, namely cluster analysis, supervised learning and dimension reduction (*Ripley, 1996, Strang, 1986, Duda & Hart, 1973*).

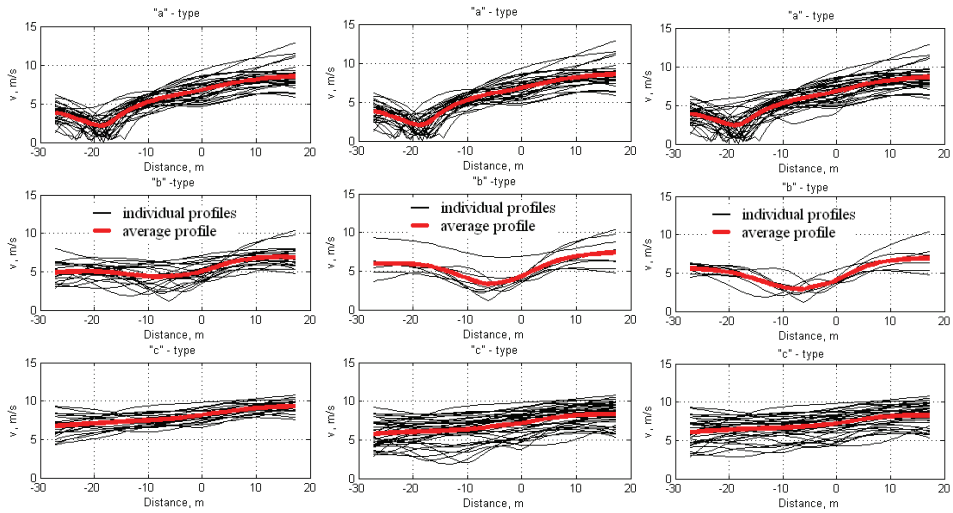
Cluster analysis is a general name for methods of dividing the data into several partitions (clusters) according to some properties considered common for the items within the cluster. Most often this property is proximity, i.e., the items in a cluster are closer to each other or to the cluster centre than to other items or other cluster centres (cluster centre in this case is also a profile with a certain shape considered "typical" by the algorithm). A clustering algorithm may force the data into a pre-defined number of clusters *k* (*k*-clustering) or find the optimal number of clusters based on the data.

The main difference in supervised learning compared to clustering is that the classification function is learnt from a training dataset containing both the input

objects and the desired outputs. The training dataset has to be produced manually beforehand. The decision is made based on the analysis of "similarity" of the classified items to each group in the training dataset.

Dimension reduction is a way to decrease the number of data points that describe each profile, but still preserve the most important information about them. This simplifies the later classification and allows visualisation of the data so that possible patterns can be seen (in this case the number of dimensions has to be reduced to less than three). In this test I use singular value decomposition technique to find the most important features in the data and represent each profile by only two co-ordinates.

Figure 26 illustrates an example of speed profile classification using these techniques.



a) *b)* *c)*

Figure 27. Classification of speed profiles by three pattern recognition techniques: a) cluster analysis (k-means); b) supervised learning (nearest neighbour); c) dimension reduction (singular value decomposition).

The general conclusion is that the pattern recognition techniques perform quite well in classifying the behaviour types, even though some variation in accuracy between the techniques can be found. The great advantage of these techniques is the automation of the classification process which allows analysis of larger datasets. Another aspect is the reduction of the subjective effects a specific observer might have on the results when doing the classification manually.

Finding the right technique for the data is often stated to be more of an art than a science, and parameters working well for one dataset may not work for another. The best strategy in this case is to have a toolbox of different techniques where the right one is found by using trials.

Profiles with shapes that do not match any of the typical patterns is a problem that needs special investigation. All three techniques are quite insensitive to such outliers and simply force them into one of the typical groups. However, examination of the outliers might be important in case they represent some kind of breakdown in normal

traffic that might have implications for safety or efficiency. Detailed examination of such situations might give an idea of how they may be eliminated. A possible solution is to compare individual profiles with the average profile and select significantly different ones.

In some cases a subjective component introduced by an observer when making classifications might be useful, especially if the differences in behaviour are difficult to express in objective terms. An observer might be able to classify quite complex traffic situations (for example, traffic conflicts) without being able to explicitly formulate the classification criteria. The pattern recognition techniques might help reveal the relations between these subjective judgements of human observers and the objective variables and contribute to a better standardisation of the conflict classification. This, however, requires a large set of traffic conflicts with detailed data on the road users' movements.

5. PRACTICAL EXPERIENCE

One of the set objectives for this thesis is to apply the video analysis system developed at Lund University to practical studies of road-user behaviour and test the system performance quality. The term "quality" is many-sided and its description should reflect applicability of the system as a measurement tool, the accuracy of the measurements, and the general usability of the system. Applicability refers to what type of studies may be done using the system and if the variables of interest can be measured using the video analysis technology. This issue has been generally discussed in the section 2.2, where one of the main conclusions is that a majority of the indicators used in behavioural studies can be calculated from the position and speed of road users extracted from a video.

Accuracy of the measurements is another important quality that determines whether the system will be used or not. If the video analysis system allows the measurement of a certain indicator "in theory", but fails to meet the requirements for accuracy, it becomes quite useless. The list of parameters characterising the accuracy depends on what task the system performs. For example, for simple detection of road users or detection of certain manoeuvres, the most important characteristics are the detection rate and the false positive rate. The detection rate shows the share of all the road users that has been correctly detected and is calculated as a ratio of the number of correct detections and some kind of "ground truth". The false positive rate shows the share of the incorrect detections (false positives) among all the detections. For detailed description of interactions, the accuracy of the position and speed of road users and, consequently, the accuracy of more complex indicators, calculated from position and speed, are important. The detection of specific situations (e.g. traffic conflicts) is again characterised by detection and false positive rates, but this time referring to the situations of interest.

As mentioned earlier (e.g. *Ismail et al., 2009*), comparison of the different systems is quite complicated since there are no defined "standard" conditions in which the tests can be done. The accuracy varies from site to site, depending on the unique combination of traffic situation and factors like camera position, viewing angle, distance to the road users, etc. at each site. For this reason it is important that the conditions in which a test has been done are also reported.

As for usability, it deals with the convenience of using the system in all the stages. This includes, for example, installation of the equipment in the field, transferring of the recorded data, interface of the software for video processing and following traffic interpretation, etc. Since the system in Lund still exists only as a prototype, many of

the supporting tasks (e.g. adjustment of settings for cameras, start of the calculations, presentation of the results) have not been automated yet. However, some general experience with the use of cameras, data management, etc., can now be reported.

The video analysis system in Lund has been used in two large-scale behavioural studies to detect cyclists moving in certain directions. In presenting the results in the following sections I concentrate on the system performance (other results from these studies can be found in Papers IV and V). I also describe a special study designed to test the accuracy of the speed and position estimates produced by different video processing algorithms. Finally, I discuss the factors that affect the system performance and conclude with the practical lessons learnt from the use of the system.

5.1. Study I – Cyclists on one-way streets in Stockholm

5.1.1. Background

Stockholm city is considering an extension of the available bicycle network by allowing cycling against traffic on some one-way streets. Hence, it was necessary to collect a large sample of observations of such cyclists in order to obtain a perspective on their "typical" behaviour, as well as frequency and types of unusual situations and traffic conflicts before the change in legislation.

5.1.2. Study design

Even though it is not legally allowed, cycling against the traffic in Stockholm is not unusual, but the frequency of such cyclists is quite low (some cyclists per hour). It is very inefficient to use human observers in such conditions as a very long time has to be spent on the spot in order to collect a sufficient number of observations. Instead, it has been decided to use video recording and then detect the "wrong-way" cyclists in the video using the automated video analysis system.

Initially, 32 sites were selected as potentially interesting for observations. However, finding good spots for the camera installations turned out to be a problem. Finally, only 22 sites were filmed of which 18 were further analysed. Three of the excluded sites did not have any one-way streets entering or exiting the intersection (only selected for controlling for general changes in cyclist flow), and the fourth was excluded since the camera turned out to be too far away from the intersection to allow proper analysis.

Eight cameras were moved between sites just before or after the weekend, resulting in three to four workdays of recording at each site. Further, the video material was processed and the objects moving in the "wrong" direction were detected with the "advanced road user detection" algorithm. Some work was done manually to ensure the quality and validate the work of the video analysis system. This included: a) calculation of the vehicle, pedestrian and cyclist flows for short periods at each site; b) visual control and sorting of the system detections, detection of the situations which might be potential conflicts.

5.1.3. Results

The recording at 18 sites resulted in 2.5 Tb of video data and 900 hours of daytime video material. After the detection performed by the video analysis system, this was reduced to approximately 27000 short video clips with a total length of 115 hours. Two observers looked through the video clips and sorted them into 4 categories: cyclists, pedestrians, cars and other (errors in video processing or odd situations). The results are presented in Table 1. The observational periods were not the same at each site; therefore, the numbers are given as an average per day.

Table 1. The results of manual classification of the automated detections at each site (average per day).

Site	Cyclists	False positives			False positive rate	Detections, total
		Pedestrians	Cars	Other		
2	147	894	11	12	86%	1063
4	100	44	19	7	41%	170
5	110	54	9	14	41%	187
6	63	938	26	126	95%	1153
7	8	13	4	1	69%	26
9	42	367	4	159	93%	572
11	35	104	29	63	85%	230
12	13	312	5	5	96%	334
14	31	140	16	48	87%	235
15	35	426	12	26	93%	497
16	55	667	7	17	93%	745
23	208	347	35	54	68%	645
27	52	163	61	61	85%	337
29	13	11	4	14	69%	42
33	55	50	9	6	54%	120
34	28	491	11	18	95%	548
36	30	15	20	10	61%	74
37	12	1	3	4	42%	19
Total	**1037**	**5037**	**285**	**645**	**85%**	**6997**

To estimate the accuracy of the automated cyclist detection, the manual counts were performed at each site for one or two 0.5-hour periods and compared with what was detected automatically during the same periods. Initially, the manual counts were expected to provide the "ground truth", but it turned out that at some sites the observers missed a few cyclists found by the automated system. Therefore, the results of manual counts were adjusted to include these cyclists, too. Table 2 presents this comparison.

Table 2. Comparison of the automatically detected "wrong-way" cyclists with the "ground truth".

Site	*Cyclists, "ground truth"*	*Automated video analysis*		
		Cyclists	*False positives*	*Detections, total*
2	*3*	*3*	*9*	*12*
4	*7*	*6*	*2*	*8*
5	*6*	*6*	*5*	*11*
6	*4*	*2*	*47*	*49*
7	*3*	*0*	*1*	*1*
9	*9*	*5*	*11*	*16*
11	*1*	*1*	*1*	*2*
12	*0*	*0*	*2*	*2*
14	*0*	*0*	*3*	*3*
15	*1*	*1*	*9*	*10*
16	*7*	*1*	*28*	*29*
23	*19*	*18*	*36*	*54*
27	*1*	*1*	*13*	*14*
29	*0*	*0*	*0*	*0*
33	*12*	*9*	*4*	*13*
34	*6*	*4*	*34*	*38*
36	*4*	*3*	*3*	*6*
37	*3*	*0*	*0*	*0*
Total	***86***	***60***	***208***	***268***

Average detection rate: 60/86=70%.
Average false positive rate: 208/268=73%.

Among all the automated detections, the observers found 43 situations that looked like potential traffic conflicts with "wrong-way" cyclists involved. However, none of these were classified as serious conflicts according to the definition used by the Swedish Traffic Conflicts Technique (*Hydén, 1987*). A small test on how well potential conflicts can be detected automatically from the video data was also performed. Site 33 was chosen for this test as it had a relatively high number of potential conflicts (6) concentrated during four 0.5-hour periods (i.e. totally 2 hours of video). The trajectories and speed profiles were extracted for all the road users in the video sequences ("trajectory extraction I" algorithm was used). Since it was known that the position was estimated with a systematic error (due to the assumption of "flat" road users) and that there were no serious conflicts to be found, the conflict criteria were set quite loosely: first, all the detected cyclist moving in the "wrong" direction were selected and then checked for encounters with other road users with

TTC < 2 sec. or TAdv < 1 sec. Table 3 compares the detection of cyclists and potential conflicts and shows the results of this test. The entire video was also watched through by an observer to get the actual number of "wrong-way" cyclists (the "ground truth").

Table 3. Detection of "wrong-way" cyclists and potential traffic conflicts by two techniques.

Video sequence	*Cyclists, "ground truth"*	*Video processing algorithm*			
		"advanced road user detection"		*"trajectory extraction I"*	
		Cyclists	*Conflicts, detected manually*	*Cyclists*	*Conflicts*
1	*9*	*8*	*2*	*9*	*2*
2	*4*	*4*	*1*	*3*	*0*
3	*3*	*2*	*2*	*3*	*2*
4	*3*	*3*	*1*	*2*	*0*
Total	***19***	***17***	***6***	***17***	***4***

Only 4 of the 6 known potential conflicts were detected automatically. Analysis of the "misses" showed that in both cases the reason was that the cyclists involved in the conflicts were not detected at all. However, the general detection rate of both techniques is quite the same (17 cyclists in both cases, but not exactly the same ones), so it might be just a coincidence that the missed cyclists were involved in conflicts.

The studied site was in the shade of a large tree for most of the day. This resulted in many false trajectories located on the shade border (as the leaves moved in the wind, the shadows were detected as separate objects). These tracks were, however, very easy to sort out as they were abnormally long timewise while the travel length did not exceed 1-2 meters.

5.2. Study II – Cyclists in roundabouts, 2 design solutions

5.2.1. Background

There are several known design options for dealing with cyclists in roundabouts, such as painted cycle lanes, separated cycle crossings and no cycle facilities. However, some differences in the safety of these solutions have been found, and the mechanisms that make one design perform better than another are still very unclear. For this study it was necessary to collect observations of cyclists at roundabouts of two design types (separated and integrated), the analysis of which would provide a better understanding of how cyclists' and motor-vehicle drivers' behaviour and interactions relate to cyclists' safety.

5.2.2. Study design

Two roundabouts with similar traffic flows and motor-vehicle speeds but different design solutions were selected for this study. In the separated roundabout (Figure 28*a*) the cycle paths, together with the pedestrian paths, run parallel to and outside the carriageways. Contact between cyclists and motor-vehicle traffic occurs only when a cyclist has to cross the carriageway at a roundabout approach or exit, interacting then with drivers entering or leaving the roundabout. In the integrated roundabout (Figure 28*b*) the cycle paths are separated from motor-vehicle traffic along the approach to the roundabout, but cyclists are led onto the carriageway and merged with motor vehicles approximately 30m before the roundabout. The intention of the design is for cyclists and motor vehicles to form one mixed flow and enter the roundabout and circulate in it as if it was just one lane.

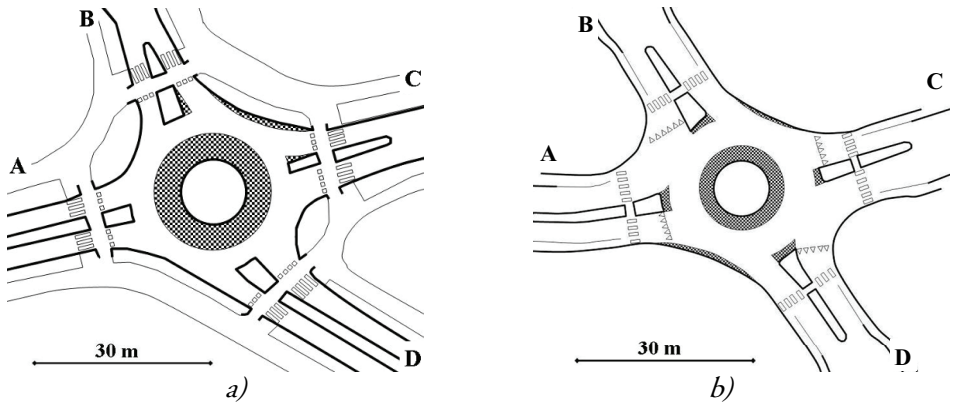


Figure 28. Two studied roundabouts: a) separated; b) integrated.

The study included both manual observations in the field and video recordings (5 days at each site). Two types of analysis were performed on the video recorded at each site:

a) a 24-hour period was watched through by observers who detected all the cyclists and classified them by travel direction, chosen path, type of interaction with motor-vehicle drivers, etc. This provided sufficient data for describing typical cyclist behaviour in "normal" situations, and the "ground truth" to which the results of automated detection were compared;

b) the automated video analysis system processed the daytime parts of the recorded video and detected the cyclists in it. These detections were looked through manually and potential conflict situations were selected.

The detection of cyclists was performed based on the size of road users in pixels and their travel paths. The trajectories of all road users were produced by a "trajectory extraction I" algorithm and then those passing through the pre-defined gates were selected (gate locations are shown in Figure 29). The size threshold was found experimentally and optimised to get the best distinction between cyclists and motor-vehicles.

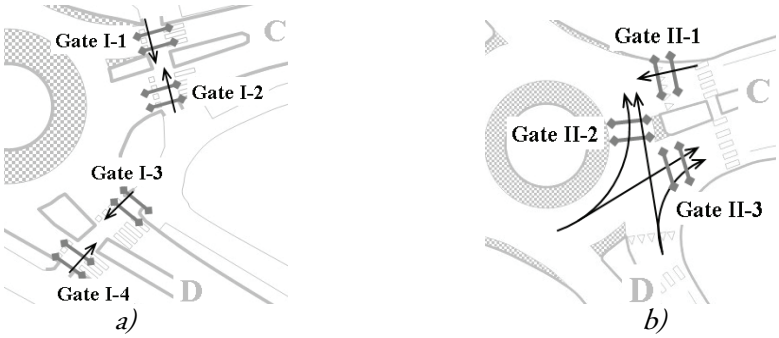


Figure 29. The schematic location of the gates used to detect cyclists at separated (a) and integrated (b) roundabouts.

5.2.3. Results

The use of the automated video analysis system "condensed" the video data from the 90 daytime hours that had been recorded down to 35 hours that were then watched by the observers. However, some cyclists were missed by the system, while some of the detections were not cyclists (false positives). Table 4 presents the comparison of the results of the automated detection and the manual detection (for 9 daytime hours from the 24-hour period that was watched through by the observers).

Table 4. Comparison between the automated and manual detections (9 hours period).

		Cyclists, "ground truth"	*Automated video analysis*				
			Cyclists	*Detection rate*	*False positives*	*False positive rate*	*Detections, total*
Separated roundabout	*Gate I-1*	149	98	66%	52	35%	150
	Gate I-2	242	104	43%	48	32%	152
	Gate I-3	99	31	31%	20	39%	51
	Gate I-4	177	32	18%	8	20%	40
Integrated roundabout	*Gate II-1*	541	387	72%	198	34%	585
	Gate II-2	172	82	48%	709	90%	791
	Gate II-3	832	485	58%	893	65%	1378

The distribution of the automated detections by type over the entire period (5 days) is shown in Table 5.

Table 5. Distribution of the automated detections by type (5 days, 45-hour period)

		Cyclists	False positives	False positive rate	Detections, total
Separated roundabout	Gate I-1	518	326	39%	844
	Gate I-2	565	328	37%	893
	Gate I-3	123	147	54%	270
	Gate I-4	131	61	32%	192
Integrated roundabout	Gate II-1	1507	556	27%	2063
	Gate II-2	381	2223	85%	2604
	Gate II-3	not available			5369

5.3. Accuracy tests

A special field study was performed to estimate the accuracy of the position and speed of road users produced by different video processing algorithms. Four cameras were used to film an urban intersection from different angles. One of the cameras was located on top of a water tower (height ≈ 25 meters and distance from the intersection ≈ 120 meters), one in a window of a nearby house (fourth floor, ≈ 12 meters above the ground) and two cameras were installed at a relatively low height (3-4 meters) on the lamppost at the intersection The camera on the water tower had a view over the entire intersection, while the other cameras covered only some parts of it.

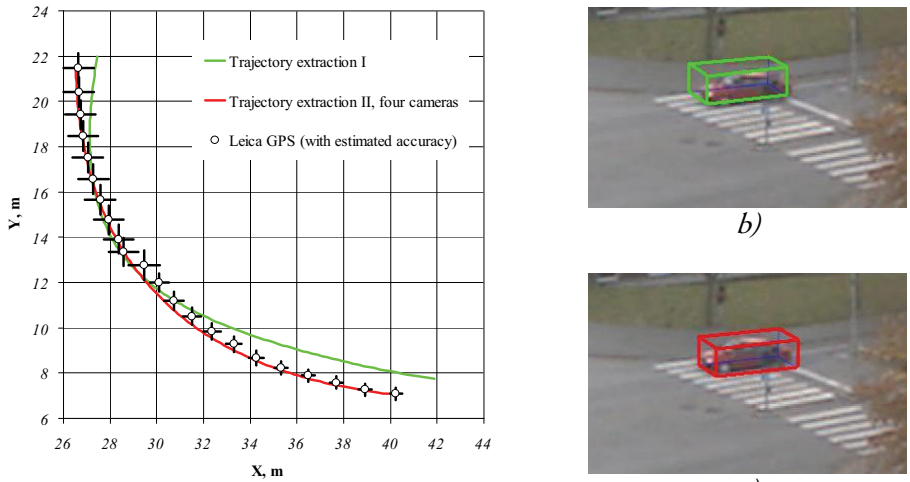
At the same time we used a car equipped with a speed logger (based on a wheel rotation counter and producing speed values with 0.5-second intervals) and a high-precision GNSS receiver[1]. In ideal conditions the GNSS receiver provides position with an accuracy of 10-20 mm, but factors like occluded satellites, multipath, ionospheric conditions, etc., may decrease the accuracy significantly. There were several buildings and some trees very close to the intersection and the actual estimated accuracy in the test varied from 5 to 70 cm (rms).

While the intersection was being filmed, the car drove through it about 20 times, coming from different approaches and performing various manoeuvres (turning right, left and driving straight ahead). Situations with intensive braking and acceleration were also simulated a few times.

Two video processing algorithms for trajectory extraction were applied to the video data. "Trajectory extraction I" used the video recorded from the water tower as it provided the best view over the intersection. For "trajectory extraction II" videos from all four cameras were used.

[1] Leica GX1230 GG, http://www.leica-geosystems.com/.

Figure 30*a* shows an example of the car's trajectory estimated by trajectory extractions I (green) and II (red) algorithms compared to the GNSS measurements (circles). The estimated accuracy of the GNSS position is also shown. Figure 30*b* and Figure 30*c* show a single frame with a simple box-model of a car projected onto the image and located according to the co-ordinates produced by the two algorithms. Figure 30*b* clearly shows how the "box" is shifted from the car's actual position ("trajectory extraction I"), while the position obtained from "trajectory extraction II" appears to be quite reasonable (Figure 30*c*).



a) b) c)

Figure 30. Position estimates by two algorithms (circles – "ground truth" from Leica GNSS, each cross represents the estimated accuracy, green – trajectory extraction I, red – trajectory extraction II): a) trajectories; b), c) vehicle box model projected onto a single frame.

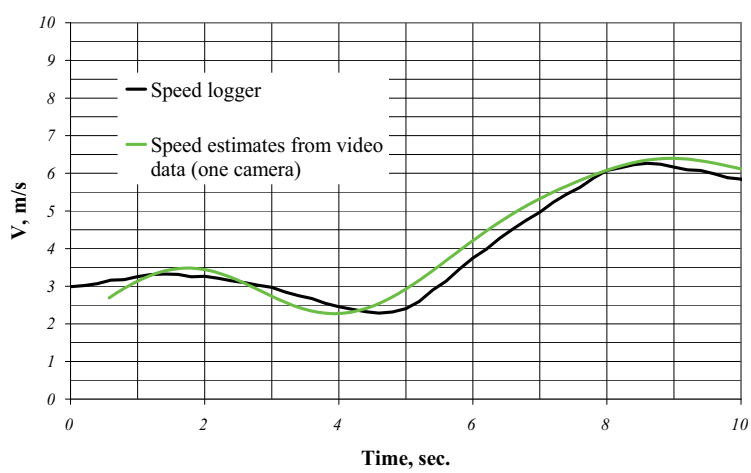


Figure 31. Speed estimates from the video data (one camera).

Figure 31 compares the speed produced by the "speed estimation" algorithm (video from the water tower was used) with the speed log inside the car. The shape of the profile estimated from a video does not exactly follow the speed log, but the difference in the speed values does not exceed 0.5 m/s, which can be considered as quite satisfactory accuracy. Possible problems might arise when the speed profile is used for calculation of the acceleration, for example, to estimate the start of an evasive manoeuvre (braking or acceleration).

5.4. Factors that affect the accuracy of the measurements taken from video

To understand the performance of the video analysis system in the studies presented here and the results of the accuracy tests, it is necessary to realise what the main underlying factors that affect the quality of the performance are. As shown in the conceptual scheme of the system (Figure 6), there are three main steps in the system operation – video data collection, video processing and interpretation of the video processing results in traffic-related terms. Obviously, what happens in each of these steps has an effect on the quality of the final output. The errors are most often accumulating and the earlier a problem occurs in this chain, the more difficult it is to compensate for it later. For instance, if the recorded video is "blinded" by sun glare at some points in time, not much can be done except try to make recordings again, this time with a better camera position and angle. However, if the problems arise during the video processing stage (e.g. many detections that are not road users, broken trajectories, etc.), it is still possible to test some other algorithms or try to filter out the "noise" during the traffic interpretation of the results. This can be done, for example, by defining some additional criteria like what the reasonable locations for correct trajectories are, how much time is reasonable for a road user to pass through the scene, etc. The broken trajectories can be connected manually if their number is not very large.

This section presents a summary of the practical lessons learnt during the work on the system and discusses the factors affecting the quality of the output and their importance. These factors are grouped around each of the steps of the system operation, i.e., into three main categories:

- quality of the input digital video recordings;
- performance of video processing algorithms that extract the traffic-related data;
- performance of the algorithms that interpret the traffic-related data.

5.4.1. Video recording quality

Camera location

The view of the traffic scene from the point where the camera is installed is crucial. The road user that is not seen cannot be detected, no matter how good the video analysis technique. The ideal location for a camera is straight above the scene, at a height sufficient to see the entire scene at a time and with no obstacles like trees,

houses, etc. around. However, the real conditions are very seldom like that. Among "trouble-makers" when choosing a spot for a camera are factors like:

- ***Furniture for camera installation.*** The camera needs to be put on something – a lamppost, balcony of a house, a flyover, etc. If there are no high and rigid objects near the site, some mobile solution (a mast) has to be used. This, however, creates problems of interference with the normal routines in traffic because of the distraction of a large new object and the evidence of video-observations going on. Some urban locations might be too cramped for mast installation, while in rural conditions the ground surrounding the road might be steep, covered with vegetation or simply not stable enough. Besides, mobile masts are less rigid than fixed installations, and since the camera "sways" in windy weather, the video has to be stabilised before the video analysis algorithms can be applied.
- ***Power supply.*** Short recordings may be made using batteries. However, for longer periods, access to an external power supply is necessary. This is not always available, especially in the rural areas. An option is to locate the batteries in an easily accessible place (e.g. on the ground) and continuously change them, but this may be quite time consuming and lead to security concerns (sabotage when left unattended).
- ***Remote control and data collection.*** If recording is done over a long period, it is reasonable to control the camera and download the data remotely. While remote control can be simple when using, for example, a mobile telephone net, downloading large amounts of raw video demands high bandwidth and requires, most often, wired connection to a computer network.
- ***Additional equipment.*** In some cases it is reasonable to have cameras connected to additional equipment (e.g. a computer) that can partly process the video data on the spot. This might save some time (analysis starts while the data is still being collected) but also reduces the amount of data to be transferred from the site. In any case, this equipment needs to be weather-protected and secured against theft or vandalism.
- ***Accessibility.*** Even if a spot for camera installation is found, it is necessary to get permission to access it. This is not normally a problem if it involves road furniture (owned by municipality or road administration), but if it is a private house it might be difficult to contact the owner or he/she might simply not want to co-operate.

Image resolution and scale

A digital image consists of pixels, small elements each having just one colour/intensity value. Image resolution is a size of the image in pixels, i.e., how many pixels it contains. The higher the resolution, the more the data available for analysis. This is double-edged – while the detection results are usually more stable for high-resolution video, the time necessary for analysis goes up dramatically with the increase in resolution.

Scale is another important quality which says how big an area in the real world is covered by one pixel. This value is usually not constant as more remote objects seem to be smaller in the image and thus have a smaller pixel size. If the number of pixels representing an object is too low, it may be classified as "noise" and therefore go undetected. This becomes crucial when an object is partly occluded and is given an even smaller pixel size. The scale can be increased by moving the camera closer to the scene, but this will decrease the total area covered by the image. Another way is to increase the resolution of the images, but this also means more data to store and analyse.

Frame rate and exposure time

A digital video is a sequence of images taken one by one with a high frequency. The frame rate characterises how many images (frames) are recorded during a time unit (usually 1 second). Again, a higher frame rate means more data and ensures more stable detections. If the frame rate is low, an object with a high speed might cover a significant distance and tracking errors, like losing or detecting new objects in the middle of the scene or mixing up 2 objects, become more probable.

The maximum frame rate is limited by the exposure time required for one frame to be recorded. When the ambient lighting is low (in twilight or night conditions), the required exposure time is longer and the frame rate is normally set to be lower compared to what is possible in the daytime.

Colour vs. grey-scaled imagery

Modern camera equipment most often records video in colour. Colour is an important quality if the recordings are meant to be visually controlled by an observer, since it creates a better representation of the scene. However, it appears that video analysis performs quite well on grey-scale video and the use of extra information provided by colour does not improve the quality of the results much, at least for the traffic scenes where the grey palette is dominating.

Atmospheric and lighting conditions

Atmospheric conditions like rain, snow or mist have a considerable effect on the quality of the recorded video as they usually make the image less clear and thus the detection more difficult. When installing the camera, it is also important to consider how the light will change during the day. Sun glare in the morning and evening can make the video data completely unusable, while the low contrast in dark shadows may affect the quality of the detections. Analysis of video recordings done in the dark requires modification of the algorithms as road users are then often poorly lighted, while the light patches on the asphalt in front of vehicles with headlamps on may be detected as moving objects.

5.4.2. Video processing algorithms

Detection and tracking of road users

Video analysis algorithms require quite a few parameters to be set, e.g. the thresholds between detection of a road user and a noise, size parameters for separation of

individual road users and their classification by type, etc. Being set for some conditions, the same parameters might no longer be optimal if the conditions change.

When a road user is occluded by some element of road furniture or another road user, most of the tracking algorithms break the trajectory and then start a new one when the road user is seen again. Some techniques (for example, the Kalman filter) allow the separate pieces of trajectories to be connected, but it is still not unusual for the same road user to be lost and detected again, each time with a new identity.

Accuracy of the position and speed estimates

The accuracy of position and speed estimates is a very important issue for studying interactions between road users. As discussed in chapter 4, a description of an interaction requires complex indicators that are seldom measurable directly, but are calculated from the speed and position of one or both road users. The inaccuracy in the "raw" data may have a substantial effect on the values of the indicators calculated from it. This is illustrated in an example in Figure 32, which shows calculations of Time-to-Collision for an encounter between a pedestrian and a car. Calculations are performed for the "true" pedestrian trajectory (*a* in Figure 32*a*) and several trajectories with introduced errors in positions generated by shifting the "true" trajectory by 1 and 2 meters (*b-e*, Figure 32*a*). Figure 32*b* shows the Time-to-Collision curves calculated for the different pedestrian trajectories. The error in position has an effect on the TTC values and how long the road users are considered to be on a collision course.

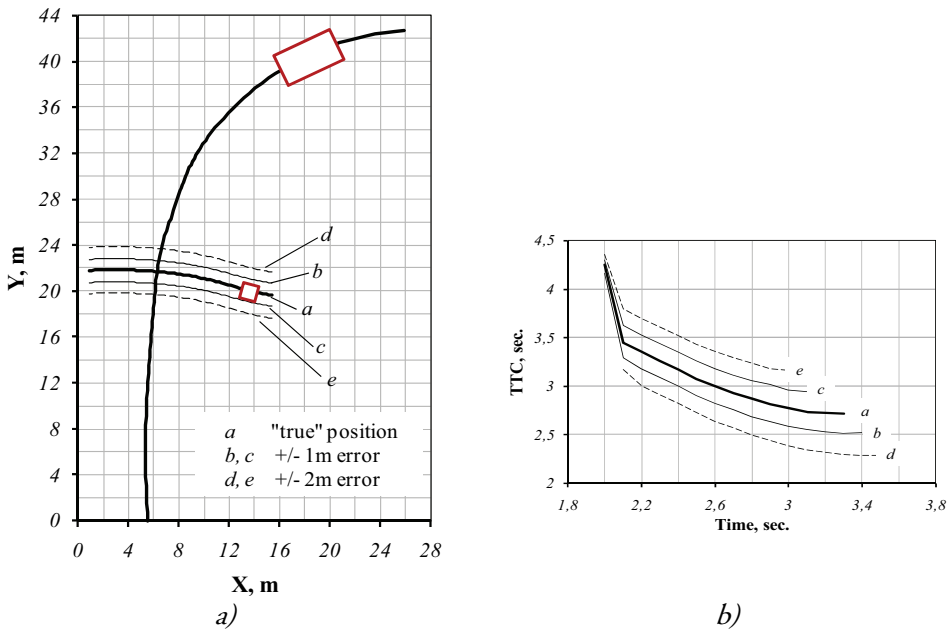


Figure 32. The effects of position accuracy on Time-to-Collision (TTC) estimates: a) vehicle and pedestrian trajectories; b) TTC profiles calculated for different pedestrian trajectories.

Getting a correct position requires estimation of the "footprint" of a road user on the ground. This in turn requires restoration of the road user's 3-dimensional shape from its 2-dimensional representation in the video images, which is not possible without making certain assumptions. The simplest assumption is that road users are "flat" and lie on the road plane. In this case, the position of a road user is estimated simply as the middle point of the pixels representing the road user in the image, and it is transferred to the real-world co-ordinates as shown in Figure 5. However, the cost of the simplicity is the introduction of a systematic error in position (Figure 33). The size of the error depends on factors like camera height above the scene, height and orientation of the road user and distance from the camera. Generally, the error is greater for large vehicles (e.g. buses, lorries) than for small ones, and increases as the road user moves away from the camera, i.e., it is not constant even for the same road user during a passage. Nonetheless, the error size does not change significantly between two adjacent frames and therefore the estimates of speed are not affected much.

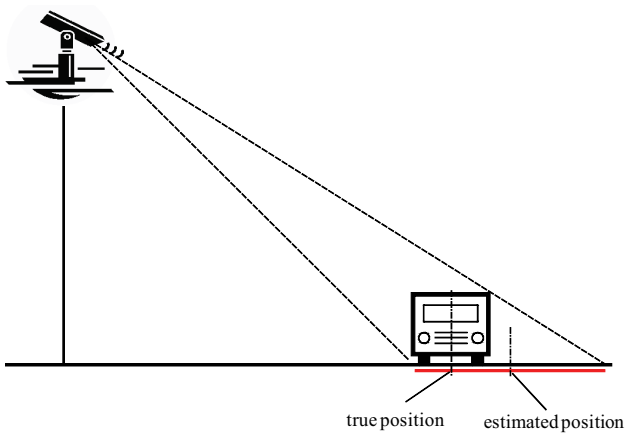


Figure 33. The systematic error introduced by assumption of "flat" road users.

Table 6. The errors in position estimation (m) caused by "flat" road user assumption (seen strictly from the side).

	Camera distance, m	***Camera height, m***		
		15	20	25
Car *width 2.2 height 1.6m*	*20*	*1.3*	*0.9*	*0.7*
	40	*2.5*	*1.8*	*1.4*
	60	*3.6*	*2.7*	*2.1*
Bus *width 2.7m height 2.5m*	*20*	*2.1*	*1.5*	*1.2*
	40	*4.1*	*3.0*	*2.3*
	60	*6.1*	*4.4*	*3.4*

Table 6 shows the results of some simple calculations of the error size for a car and a bus position depending on the camera height and the distance to the vehicle. Here, it

is assumed that the vehicle is seen strictly from the side, i.e., the length parameter does not have any effect on the error size.

Algorithms that use some pre-defined shapes of road users avoid this type of error. However, some inaccuracies are still possible if the real size of the road user differs from the size used in the model.

5.4.3. Traffic data interpretation algorithms

The traffic-related data produced by video analysis algorithms may already contain some errors like missed road users or false detections, trajectories that are split into several pieces or systematically shifted from the true position, "swapped" identities of road users, etc. However, it is possible to compensate for some of these errors in a later stage. For example, if selection of certain types of manoeuvres is done by setting some "gates" that the trajectory should pass through, the position of the "gates" may be chosen in areas where no splits in trajectories are expected and the gates' positions can be shifted in the same direction as the trajectories. The final gate position might appear quite far away from its "intuitive" position on the road and some experimenting is required to find a location that provides good detection results.

5.4.4. Accuracy and observation period

Choosing a video analysis technique to be used in a particular application is always a trade-off between accuracy of the detections and position estimation, on the one hand, and the computation intensity, on the other, as higher accuracy usually requires more sophisticated algorithms which also take a longer time to run. For example, in applications like traffic counting or detection of congestion the position does not have to be very accurate, but the results are much more valuable if delivered in real-time or close to it. On the other hand, detailed analysis of interactions between road users requires very precise position and speed description, but can often be performed off-line.

Another important factor to consider is how long an observation period is necessary to collect a sufficient amount of the traffic-related data. A sample of relatively frequent events, for example vehicles making a left turn at an intersection, can be collected in several hours. Other types of events (e.g. traffic conflicts or accidents) can be quite rare and therefore the observations have to be carried out for much longer (weeks or months). However, the length of the observation period is also limited by the time required to process the recorded video, and in some cases it might be necessary to choose a simpler analysis technique to make the waiting time for the results more reasonable.

6. DISCUSSION

6.1. The video analysis system - current state and future development

When we started the work on development of the automated video analysis system at Lund University, we had ambitions to deliver a tool that would provide detailed description of road users' motion sufficient for making studies of their behaviour, allow analysis of long observation periods and, as an ultimate task, automatically detect and classify the critical incidents in a traffic process – the traffic conflicts. When this thesis is completed, we will only be half-way to reaching the set goals. The existing version of the system may thus be seen as a working prototype to make simple studies and tests of usability, as well as a stimulus and test ground for further development of the theories and methods for relating individual behaviour to the important qualities of a traffic system.

In its current version, the system has quite a few limitations. The results of Studies I-II show that the accuracy of road user detection varied between the studied sites. There was also great variation between parts of the scene even at the same site. In Study I ("wrong-way" cyclists), the intensive pedestrian flows resulted in many false positives, as the system was unable to distinguish between cyclists and pedestrians properly. This indicates that filtering algorithms, not just based on road-user size or number of interest points, but rather analysing the shape of the road users, are necessary. However, the video analysis performed better than human observers in some cases. This happened in very crowded conditions with lots of pedestrians and cyclists mixed and moving in different directions, which probably distracted the observers a lot.

In Study II (cyclist in roundabouts), several cyclist flows were studied with video from only one camera at each site. It turned out to be very difficult to find the optimal view for all the cyclist directions, and it was necessary to prioritise to get a better view of at least some of the flows. This is clearly seen in much better detection rates at the gates for which the camera was optimised (gates II-1 and I-2 at the separated roundabout and gate I-1 at the integrated one). Generally, if such a compromise is to be made, it is important to have an idea about what directions are more important for studying, which, in its turn, might require some pilot observations to be performed before the installation of the cameras.

There are some indications that accuracy of detection and tracking depends on traffic conditions. Road users often occlude each other in dense traffic, and as a result are

lost by the tracker or swap identities. An important task is to study the relationship between the quality of video analysis results and traffic parameters, as this directly reflects the reliability of the technique. As long as the road users are "lost" unsystematically, the misses may be partly compensated for by simply increasing the observation periods and thus the sample size. This will not provide the absolute numbers correctly (e.g. the total number of road users performing a certain manoeuvre), but at least enables to get reliable relative values (e.g. relative frequency of certain manoeuvre types). However, if the studied phenomenon and the detection accuracy depend on the same factors, there is a risk of introducing systematic bias in the results.

The small conflict detection test performed in Study I ("wrong-way" cyclists) does not allow us to make any solid conclusions since it is based on very limited conflict data. What is obvious is that a detailed study of interactions and detection of possible conflicts requires quite accurate estimation of road users' position, size and speed. On the other hand, the most interesting events in traffic (at least from a safety perspective) are also very rare, which means that the observation period has to be quite long to allow observation of a sufficient number of such events (e.g. a typical conflict study takes at least 3-5 days). Of the two trajectory extraction algorithms available, only algorithm I is "quick" enough to allow analysis of longer video sequences in reasonable time. However, the position accuracy it provides is not sufficient to reliably calculate safety indicators like Time-to-Collision, Time Advantage, etc. Position errors become critical in situations when road users pass each other with small (but perfectly safe) margins, for example on a parallel course in two adjacent lanes when they are very often detected as having collided.

Trajectory extraction algorithm II appears to provide much more accurate positions. The problem, though, is that it requires very long computation time and therefore cannot be used for analysing long video sequences. A possible compromise between the need for accuracy and limits to calculation time is to analyse the video data in two steps, first detecting potentially relevant situations with very simple and fast algorithms, and then analysing these detections once again with more accurate algorithms that require longer calculation time. After the second step, since a lot of uninteresting video is removed, it is also possible to use human observers to look through the detections and classify them and/or extract information that is not possible for automated algorithms to retrieve (e.g. age and sex of road users, informal signals or other forms of communication between them, etc.).

When using "simple" detectors it is important to balance two requirements: *i*) that the threshold is low enough to ensure that rare but relevant events (e.g. serious conflicts) are included in sufficient numbers and *ii*) that the level is high enough to prevent too many false detections that make the detector useless. An example of the latter was the average false positive rate of 85% in Study I ("wrong-way" cyclists). Here the detector parameters were set too low and the separation of the correct detections from the false positives became very laborious. In some cases, if it is important to detect *all* the relevant events that are very rare (e.g. accidents), it might still be reasonable to have a very low threshold and thus a high false positive rate.

The problem of long computation time is partly mitigated if several computers are used to analyse different parts of the video at the same time, running day and night without the direct control of an operator. However, for longer observation periods (order of months or years), the extremely long waiting time to get the results greatly diminishes their value. On the other hand, the present hardware limitations are not permanent and should not block further development of the "soft" part of the technology. After being tested on today's hardware, with a smaller amount of input data and longer calculation time, the new algorithms could be implemented and used on a larger scale as soon as the proper hardware becomes available. Nonetheless, it still seems reasonable to further investigate how the programming codes can be optimised, and if the advantages of parallel programming for several processors can be utilised.

Another problem is that in many cases the view provided by one camera is not sufficient to cover the entire studied scene, and finding a good place for camera installation is always a challenge. From this perspective, it is a great advantage for a video processing algorithm to be able to integrate the data coming from several cameras, each viewing a part, but together covering the entire scene (as in the case of trajectory extraction algorithm II). It is becoming more and more common to install cameras for purposes other than traffic observations, and together they cover relatively large areas (e.g. *Conche & Tight, 2006* report that CCTV, an acronym for close-circuit television, in British cities allows video recordings of about a quarter of all traffic accidents). Integration of the data from such cameras into the system can greatly decrease the problems of getting the necessary permission, installation, access to the recoded data for downloading, etc.

An alternative way to get a better view for one camera is to use aircraft (e.g. a helicopter) that can observe large scenes at a time from a greater height (e.g. *Zhao & Nevatia, 2003*). This, however, requires additional calculations to stabilise the images taken from a moving camera. The time that a helicopter can hover in the air is quite limited and for longer observation periods some other flying alternatives that are not as energy consuming should be employed (e.g. air balloons or dirigibles). There are also quite strict regulations that limit the use of pilotless aircraft in urban areas. For the moment, there are no universally acceptable devices and the price for developing them might be too prohibitive.

Lighting and atmospheric conditions greatly affect the performance of video analysis algorithms. Studying the behaviour of the road users when visibility is less favourable is also important; hence additional improvements are necessary to make algorithms more stable when analysing data collected in such conditions. Poor-quality video data may possibly be complemented with data from some other sensors that are less light-dependent. These sensors may be used either as simple detectors indicating that a road user is present, and thus that extra attention should be paid to the video (may be done, for example, with radar or ultra-sound detectors), or as sources of high-resolution data that can be analysed in a similar manner to videos (e.g. lidars or infra-red cameras). The information provided by the additional sensors can also be useful in good visibility conditions, for example, the profiles detected by inductive loops can be used to verify the type and speed of a vehicle. The integration of several sensors of

different types poses a question of how these data can be fused (*Wender & Dietmayer, 2007*).

6.2. Behavioural studies and safety evaluation based on video data

Using video analysis to produce data has many advantages compared to other conventional instruments. Video itself is very illustrative, and the data extracted from it can cover a large population of road users and various types of situations. Other advantages are spatial and temporal linkage of the measured variables and high temporal resolution of the data that provides nearly continuous description of the road users' motions. This allows for a more advanced analysis by introducing complex indicators (for example, "sequence of values" instead of the "single value" indicators) and using methods that can take advantage of the continuity of the data.

So far, many conventional instruments that enable "single value" measurements have been superior to video analysis in accuracy and simplicity of the data collection (e.g. a radar gun delivers speed values directly and does not require special installations). The sequential data, as produced by video analysis, might be more difficult to collect, but allows analysis of the processes in traffic instead of the states at certain moments. Modifying a conventional instrument for sequential data collection is not always simple and in some cases impossible (for example, a human observer has very limited capacity when it comes to taking a sequence of measurements within a short time).

However, introduction of new indicators is to be done with some caution. An indicator should not be collected just because "it is so easy to collect now", but its validity has to be tested first. The issue of the indicators' validity gets very little attention in the publications used in the literature study and, except for very few thoroughly studied indicators (I refer here to serious conflicts and speed and their relation to the number of registered injury and fatal accidents), the relation between an indicator and the described quality is far too often based on assumptions and common sense. Establishing the validity is quite a challenge as it requires large data samples representing different locations and conditions. In this respect, video analysis can contribute by being an effective tool that can collect extensive datasets, even for studies that are not normally feasible due to high costs of data collection with conventional tools (for example, at locations with low traffic intensity where observation periods have to be very long).

On the other hand, when making a decision on what indicators to use, it is necessary to consider what indicators were used in other similar studies. So far, individual behavioural studies in traffic have been restricted to a few locations and short observation periods and seldom conducted on a large scale. Therefore, there is a need to compare the results or even merge the data from different studies, which is only possible if the indicators used in each study are defined and applied in the same way. Implementation of indicators in a video analysis system requires very strict and objective definitions and operational procedures for their calculation. The side effect of this might be that before an indicator is implemented, the questions of its validity and standard definition is raised and discussed once again.

The wider use of indirect traffic safety indicators has also been limited by the problems of establishing their validity, which requires collection of large datasets of both indirect indicators and accidents. For example, even though they are much more frequent than accidents, serious traffic conflicts are still too rare to allow collection of large samples using human observers. Although it is known (*Hydén, 1987*) that the relation between serious conflicts and injury and fatal accidents depends on the type of the conflicts (e.g. the angle of approach and types of road users involved), it is not always possible to split the available conflict data into many sub-categories as the number of conflicts (and accidents) in each group will be too few. A possible way to extend the accident dataset is to include slight injury and property-damage-only accidents that normally are not reported to the official statistics. Most probable, these accidents also correlate with the serious conflicts, but this hypothesis has never been possible to test as no data about such accidents was available. From this perspective, video analysis has a great potential to contribute to validation of the various safety indicators, since observations carried out over long periods will yield a large sample of the indicator measurements and the actual accidents, including the minor accidents. The entire process of accident development can be studied and compared to the processes of the near-accidents and normal encounters. However, to make such studies, the observation periods have to be significantly extended (order of months and years compared to the current periods of a few days).

The construction of severity hierarchies requires the definition of some universal measure of severity. This is not a simple matter since this measure has to reflect both the accident risk and severity of possible consequences, make the produced hierarchy shape as close as possible to the "true" hierarchy shape and be operational enough to be applied to all the possible encounter types (with and without a collision or crossing course, different types of road users involved, etc.). There are some indications that human observers intuitively use some subjective measure of the severity that might fit these requirements. One of the conclusions from the international calibration study of traffic-conflict techniques from different countries (*Grayson, 1984*) is that "even the observers were instructed to use specific cues such as TTC or PET, they will incorporate other aspects of the situation as well. Although severity scaling is linked to objective measures, it also includes a subjective dimension. This results in common understanding of conflict severity, at least for trained observers". *Svensson, 1992*, reports that the serious conflicts classified subjectively by observers correlate better with accidents than the conflicts defined strictly by the definition of the Swedish Traffic Conflicts Technique. The challenge, however, is to find an objective and operational measure that corresponds to the subjective severity judgements and validate it.

The advantage of using continuous indicators is that they allow us to study the development of an encounter as a process and test another approach to classifying the severity of encounters. Instead of using indicator values at a certain moment (as in the case of TTC_{min}, TA or PET), the entire indicator profiles, i.e., TTC or TAdv curves, can be examined. Such examination might reveal what typical shapes characterise "normal" and "critical" situations, and the "nearness" of a profile to one or another

"typical" shape may be used as a measure of the severity. Pattern recognition techniques could be a valuable tool in such analysis.

Some of the indicators that are important for estimation of the encounter severity, especially the severity of consequences in case of a collision, are difficult to collect using video analysis. This refers to the use of helmets by cyclists, motorcycle and moped drivers, use of safety belts by car drivers and passengers, age of road users, especially the vulnerable ones (pedestrians and cyclists), etc. In some cases, the necessary information can be extracted if a human observer looks through situations that have been selected automatically using some other criteria.

The proposed set of indicators describing an encounter is a first step towards finding a universal severity measure. Some work on validation of TTC and PET as severity measures in conflict situations has already been done (*van der Horst, 1990, Hydén, 1987, Grayson, 1984*). TTC was found to better reflect the severity than other indicators, for example, PET. It is still possible that TAdv combined with T_2 might be a better measure of severity than PET on its own. Other indicators, for example, TET, TIT or time-inversed TTC (1/TTC, *Kiefer et al., 2005*) should also be tested on a larger scale.

The definition of many of the proposed safety indicators is based on a concept of "the same paths and speed", i.e., the motion of a road user has to be projected in the future. A human observer can (in most cases) make such a projection relatively easily, but it is difficult to explain exactly how this is done. It is possible that the price for seeming "easiness" is groove simplifications done, most probably, quite unconsciously, e.g. "compression" of a conflict zone into some vague "conflict point", treatment of all approaching angles as if they were right angles, etc. When the projection is to be done automatically, clear and unambiguous algorithms are required. A simple assumption of travelling along a straight line does not work in the case of a road user making a turning manoeuvre, since in this case the potential collision point takes quite an unrealistic position. A possible approximation is to assume that road users actually follow the planned path, i.e., to use the known trajectory (if the indicator is calculated after the trajectories have been extracted). This may be misleading if the road user avoids a conflict by changing paths, for example taking a larger radius in a turn or changing lanes. Another alternative is to use an "average" path, calculated from the trajectories of many road users making the same manoeuvre. The problem, however, is that in critical situations the paths might not follow the average pattern. A detailed analysis of critical situations might reveal when the deviation from the "average" pattern starts to develop during an encounter, and if the high severity of the situation can be detected *before* that moment, i.e., when the "average" assumptions are still valid.

Different variations of TTC definitions were tested by *van der Horst, 1990*, for example based on assumptions of constant angular velocity and constant acceleration of a vehicle (this is supposed to represent a situation when a driver is no longer controlling the vehicle, and the steering wheel and the gas pedal positions are kept unchanged). The paths calculated with constant angular velocity easily take very peculiar shapes and lead outside the road. As for constant accelerations, the TTC

values are still reasonable, but there is no clear evidence that the predictive power of TTC has improved. Still, further tests on the use of acceleration in calculation of the proposed indicator set are necessary.

Another theoretical problem is the assumption that an elementary event that can result in an accident is an encounter between two road users. This totally excludes situations with only one road user involved, even though single accidents are common and, for example, contribute 32 % of all the traffic fatalities in Sweden (*SIKA, 2009*). The main difference in single accidents is that the factors contributing to the accident risk (e.g. fatigue causing a driver to fall asleep) cannot be attributed to some particular units of a road infrastructure (e.g. an intersection), but are spread over the entire network. Thus, it is not possible to study such accidents by making observations at a certain site; it has to be done, for example, from inside a vehicle. Some attempts to register traffic conflicts from vehicles are reported in the literature (e.g. *Nygård, 1999, Risser, 1985*). It is reasonable to assume that, similar to encounters, single-road-user-events belong to some kind of severity hierarchy, i.e., accidents, near-accidents when the road user manages to regain control of the vehicle at the very last moment and avoids the collision, and so on. It is possible to modify some of the indicators developed for encounter description so that they are applicable to single-road-user-events (e.g. the Time-to-Lane Crossing, TLC, is an extension of the TTC-concept and describes the time remaining for a road user to reach one of the lane boundaries – *van Winsum et al., 2000*). Still, further research is necessary on how these events may be integrated into the hierarchy based on encounters.

When a severity hierarchy is created, an important question is how it is to be interpreted. It is important to elaborate on what the whole shape and frequency of events at different levels represent. The severity hierarchies proposed earlier (*Svensson & Hydén, 2006, Svensson, 1998*) include only events with a collision course. It is argued that interactions at fairly high severities may be positive from a safety point of view because they are frequent and severe enough to increase awareness. My suggestion is to include encounters without a collision course in the severity hierarchy, too, since such encounters also have the potential to become accidents if the spatial and temporal relation between the road users changes. It will be interesting to analyse whether these extended hierarchies can also be interpreted in a similar manner to a "collision course-only" hierarchy. With information about the encounter processes and the severity of these processes, it will be possible to formulate and test hypotheses on the interrelationships of design of the traffic environment, behaviour and risk. The final goal will be to have an operational and usable tool for safety estimation, similar to today's traffic-conflict techniques, but with much higher degrees of validity, reliability and automation, which can be disseminated and used by traffic safety engineers on a daily basis.

The prospect of being able to collect data on road users over longer time periods and, possibly, over larger areas allows us to make new types of studies and observe completely new phenomena. Applied to vehicle traffic, it will be interesting to study the behaviour over longer road sections, for example adaptation of speed to the road geometry and elements of the infrastructure, interactions during lane change manoeuvres and near lane merging locations, etc. At the moment, our department is

launching several projects within the framework HASTA (Sustainable Attractive City – *HASTA, 2009*), aimed at studying the "overall transport quality" of a city where aspects like safety, health, security, accessibility, comfort, equality, participation and environment are treated integrally. The integral approach requires introduction of conceptually new indicators that still have to be developed. However, even now there is no doubt that many of the indicators will eventually be based on the micro-level behaviour of the "city users" and video analysis will be an indispensable tool for such data collection.

7. CONCLUSIONS

The main conclusion of this thesis is that automated video analysis technology has great potential for traffic-behaviour studies. The majority of micro-level behaviour indicators can be extracted from video data. The indicators that are hard to obtain from video describe road users' qualities like age, gender and actions like informal signals, eye contact, etc. and have to be collected using other methods (e.g. human observers). Compared to other conventional tools, video analysis allows for much more extensive data collection with regard to the lengths of observation periods and the level of details (e.g. continuous description of the traffic processes rather than the static states at certain moments). Video analysis also opens up for standardisation of behaviour studies both regarding data collection and analysis.

The detailed description of the processes in traffic contributes to a better understanding of the mechanisms that lie behind the development of the normal and critical situations (including accidents) in traffic. Organisation of the elementary traffic events into severity hierarchies creates a better illustration of the safety situation, and allows us to study the balance between safety and other qualities valued by road users, for example, comfort, efficiency and mobility.

The video analysis system at Lund University is built on the principle of a toolbox, i.e., it includes a number of techniques. The choice of the right technique for a particular application is always a balance between the accuracy of the results and the intensity of the computation, requirements as to the quality of the input video data and amount of work on setting the necessary parameters. Improvements of the output quality most often require the use of more advanced algorithms, and thus a longer time to produce the results. On the other hand, the constant progress and improvements of the hardware parameters allow more and more intensive calculations to be run within a reasonable time.

Using a video analysis system in two large-scale studies has shown that further improvements of accuracy are necessary. This concerns detection and tracking of the road users, distinguishing of the road user types and accuracy of the position estimation. The latter is especially important for detailed studies of the interactions between road users and calculation of the safety indicators. When the appropriate accuracy is reached, the system will provide a unique opportunity for validation of the theories relating the behaviour during individual interactions to the general safety level, and studies of the relations of the main qualities (safety, efficiency, comfort, etc.) in traffic systems.

Future work on system improvements should also include simplification of the video data collection (e.g. use of the video recorded at locations with a poor view, utilisation of existing camera installations) and extension of the area size studied by the system.

A very important future application of the system is further validation of the traffic-conflicts techniques and other theories that relate the micro-level behaviour of road users to the qualities of a traffic system.

Working in a mixed team of traffic researchers and developers of video analysis algorithms has been a very fruitful learning process. Such work stimulated the exchange of expertise and ideas, created opportunities for seeing the problems from other perspectives and, finally, contributed to better quality of the video analysis systems.

8. ACKNOWLEDGEMENT

I would very much like to thank my supervisors *Christer Hydén*, *Åse Svensson* and *Karin Brundell-Freij* for long hours discussing my thesis, the fascinating world of traffic safety science and life and its "ups and downs". I'm grateful for their support and understanding and belief in me from the very first day we met.

I want also to thank the maths-team, *Håkan Ardö*, *Kalle Åström* and *Rikard Berthilsson*, for introducing me to video processing, the "black box" in which I slowly learnt to see and understand more and more. Special thanks are due to Håkan for solving many practical camera-related problems and for being a very interesting travel-mate and a reliable co-writer and co-speaker on many occasions.

I am grateful to all my colleagues at the Traffic and Roads, Institution for Technique and Society, for encouraging me to speak Swedish, for helping with many practical problems I met when I came first to Lund, for kind smiles and deep talks. Special thanks to *András Várhelyi* – for letting me help with many of his courses and learn from that, *Mia Sinclair* & *Inger Myhrén* – for being able to solve any problem at any time, *Ralf Risser* – for being my room-mate and for the exciting ICTCT meetings, *Thomas Jonsson* – for his great interest in our system and help in finding applications for it, *Lars Ollvik* – for exciting GPS experiments, *Jutta Pauna* – for being kind and helpful, *Georg Siotis*, *Martin Veeborn* & *Johannes Eriksson* – for keeping my computer alive and my spirit high ☺.

It was a great pleasure to be a member of the *doktorandgäng*, to participate in courses together, play hockey, organise parties, make disputation films and do many other fun things. Thank you – *Emeli*, *Andreas*, *Emma*, *Robert*, *Lina*, *Lisa*, *Marlene*, *Hanna L* & *Hanna W*, *Charlotte*, *Annika*, *Jenny*, *Omar*, *Till*, *David*, *Erik*, *Vanessa* and *Nina*.

Thanks to my family and especially my wife *Helene* who took my frequent travels and work on weekends with great patience, loved and inspired me and made me a better person.

Thanks to my Belarusian support team: my parents *Julia* & *Viachaslau* – for calling me and waiting for me every summer and *Yury Vazhnik* – for inspiring me to study in Sweden, keeping in contact with me all these years and coming to Lund many times.

9. REFERENCES

Agarwal, A., B. Triggs (2006). Hyperfeatures: multilevel local coding for visual recognition. The 9th European Conference on Computer Vision (ECCV).

Allen, B. L., B. T. Shin, P. J. Cooper (1978). Analysis of traffic conflicts and collisions. Transportation Research Record 667, pp. 67-74.

Amundsen, F. N., C. Hydén (1977). Proceedings of the first international traffic conflicts technique workshop. Institute of Transport Economics, Oslo.

Andersson, J. (2000). Image processing for analysis of road user behaviour – a trajectory based solution. Lund Institute of Technology, Department of Technology and Society, Traffic Engineering. Bulletin 212.

André, M., J. Hickman, D. Hassel, R. Jourmard (1995). Driving cycles for emission measurements under European conditions. SAE. Technical Paper 950926.

Archer, J. (2005). Indicators for traffic safety assessment and prediction and their application in micro-simulation modelling: a study of urban and suburban intersections. Doctoral thesis, Royal Institute of Technology, Stockholm, Department of Infrastructure.

Ardö, H. (2009). Multi-target tracking using on-line viterbi optimisation and stochastic modelling. Doctoral thesis, Lund University, Faculty of Engineering LTH, Centre for Mathematical Science.

Aronsson, K. F. M. (2005). Speed characteristics of urban streets based on driver behaviour studies and simulation. Doctoral thesis, Royal Institute of Technology, Stockholm, Department of Transport and Economics.

Arranz, J. M., A. I. Gil (2009). Traffic accidents, deaths and alcohol consumption. Applied Economics 41 (20), pp. 2583-2595.

Asmussen, E. (Ed.) (1984). International Calibration Study of Traffic Conflict Techniques. Springer-Verlag, Berlin Heidelberg.

Atev, S., H. Arumugam, O. Masoud, R. Janardan, N. P. Papanikolopoulos (2005). A vision-based approach to collision prediction at traffic intersections. IEEE transactions on intelligent transportation systems 6 (4), pp. 416-423.

Autoscope® (2009). Video detection solutions to improve traffic flow and roadway safety. www.autoscope.com *Acc.* 9 February 2009.

Baguley, C. J. (1984). The British traffic conflicts technique *in* E. Asmussen (Ed.) International calibration study of traffic conflict techniques. Springer-Verlag, Berlin Heidelberg New York Tokyo.

Berclaz, J., F. Fleuret, P. Fua (2006). Robust people tracking with global trajectory optimization. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, USA.

Berntman, M. (2003). Consequences of traffic casualties in relation to traffic-engineering factors - an analysis in short-term and long-term perspectives. Doctoral thesis, University of Lund, Lund Institute of Technology, Department of Technology and Society.

Carsten, O. M. J., D. J. Sherborne, J. A. Rothengatter (1998). Intelligent traffic signals for pedestrians: evaluation of trials in three countries. Transportation Research Part C 6, pp. 213-229.

Coifman, B., D. Beymer, P. McLauchlan, J. Malik (1998). A real time computer vision system for vehicle tracking and traffic surveillance. Transportation Research Part C 6, pp. 271-288.

Conche, F., M. Tight (2006). Use of CCTV to determine road accident factors in urban areas. Accident Analysis & Prevention 38, pp. 1197-1207.

Cooper, P. J. (1984). Experience with traffic conflicts in Canada with emphasis on "post encroachment time" techniques *in* E. Asmussen (Ed.) International calibration study of traffic conflict techniques. Springer-Verlag, Berlin Heidelberg New York Tokyo.
Corneliu, T., S. Nedevschi (2008). Real-time pedestrian classification exploiting 2D and 3D information. Intelligent Transport Systems, IET 2 (3), pp. 201-210.
Dalal, N., B. Triggs (2005). Histograms of oriented gradients for human detection. IEEE Conference on Computer Vision and Pattern Recognition.
Duda, R. O., P. E. Hart (1973). Pattern classification and scene analysis. A Wiley-interscience publication, New York.
Ekman, L. (1996). On the treatment of flow in traffic safety analysis: a non-parametric approach applied on vulnerable road users. Department of Traffic Planning and Engineering, Lund University. Bulletin 136.
Elvik, R. (2008). Criteria for causal inference: lessons from epidemiology applied to safety evaluation. Workshop on future directions in highway crash data modelling, Transportation Research Board, National Academy of Sciences, Washington D. C.
Elvik, R., P. Christensen, A. Amundsen (2004). Speed and road accidents: an evaluation of the Power Model. TOI. report 740/2004.
Elvik, R., T. Vaa (2004). The handbook of road safety measures. Elsevier.
Englund, A., N. P. Gregersen, C. Hydén, P. Lövsund, L. Åberg (1998). Trafiksäkerhet – en kunskapsöversikt (*in Swedish*) Traffic safety - a knowledge review. KFB, Studentlitteratur, Lund.
Ericsson, E. (2000). Variability in urban driving patterns. Transportation Research Part D 5, pp. 337-354.
Ervin, R., C. MacAdam, J. Walker, S. Bogard, M. Hagan, A. Vayda, E. Anderson (2000). System for Assessment of the Vehicle Motion Environment (SAVME), vol. I. National Highway Traffic Safety Administration. Report No. UMTRI-2000-21-1.
ETSC (2001). Transport safety performance indicators. European Transport Safety Council.
Evans, L. (1986). Riak homeostasis theory and traffic accident data. Risk Analysis 6 (1), pp. 81-94.
Evans, L. (1991). Traffic Safety and the Driver. Van Nostrand Reinhold, New York.
Fildes, B. N., G. Rumbold, A. Leening (1991). Speed behaviour and drivers' attitude to speeding. Monach University, Accident Research Centre. Report No. 16.
Gettman, D., L. Head (2003). Surrogate safety measures from traffic simulation models. U. S. Department of Transportation, Federal Highway Administration. Report No. FHWA-RD-03-050.
Grayson, G. B. (1984). The Malmö study: a calibration of traffic conflict techniques. SWOV. Report R-84-12.
GRSP (2007). Drinking and driving: a road safety manual for decision makers and practitioners. Global Road Safety Partnership.
Hakkert, A., V. Gitelman, E. Ben-Shabat (2002). An evaluation of crosswalk warning systems: effects on pedestrian and vehicle behaviour. Transportation Research Part F 5, pp. 275-292.
Hansson, A. (1975). Studies in driver behaviour, with applications in traffic design and planning. Two examples (*in Swedish*). Lund Institute of Technology, Department of Traffic Planning and Engineering, University of Lund. Bulletin 11.
HASTA (2009). Hållbar Attraktiv Stad (*in Swedish*) Sustainable Attractive City. http://www.tft.lth.se/hasta/ *Acc.* 3 Dec 2009.
Hauer, E. (1997). Observational before-after studies in road safety. Pergamon.
Hayward, J. C. (1971). Near misses as a measure of safety at urban intersections. Doctoral thesis, The Pensilvania State University, Department of Civil Engineering.
Hitachi (2009). Road Transportation Image Processing Systems. http://www.hitachi.co.jp/en/products/its/product/infrastructure/2104372_45058.html *Acc.* 1 Dec 2009.
Hoogendoorn, S. P., W. Daamen, P. H. L. Bovy (2003). Extracting microscopic pedestrian characteristics from video data. Transportation Research Board (TRB), 82nd annual meeting, Washington, D. C.
Hydén, C. (1987). The development of a method for traffic safety evaluation: the Swedish traffic conflict technique. Doctoral thesis, Lund University, Department of Traffic Planning and Engineering.

Hydén, C., A. Várhelyi (2000). The effects on safety, time consumption and environment of large scale use of roundabouts in an urban area: a case study. Accident Analysis and Prevention 32, pp. 11-23.

Isard, M., J. MacCormick (2001). BraMBLe: A bayesian multiple-blob tracker. IEEE International Conference on Computer Vision.

Ismail, K., T. Sayed, N. Saunier, C. Lim (2009). Automated analysis of pedestrian-vehicle conflicts using video data. The Transportation Research Board (TRB) 88th Annual Meeting, Washington D. C., USA.

Karlgren, J. (2001). Bilars hastighet längs gator med gupp (*in Swedish*) Vehicles' speed along streets with humps. Doctoral thesis, Chalmers Tekniska Högskola, Göteborg, Tema Stad & Trafik, Sektionen för Arkitektur.

Kiefer, R. J., D. J. LeBlanc, C. A. Flannagan (2005). Developing an inverse time-to-collision crash alert timing approach based on drivers' last-second braking and steering judgments. Accident Analysis and Prevention 37, pp. 295-303.

Kirchhof, M., U. Stilla (2006). Detection of moving objects in airborne thermal videos. ISPRS Journal of Photogrammetry & Remote Sensing 61 (3-4), pp. 187-196.

Kloeden, C. N., A. J. McLean, V. M. Moore, G. Ponte (1997). Travelling Speed and the Risk of Crash Involvement. Volume I: findings. University of Adelaide, Road Accident Research Unit. Report No. CR 172.

Koller, D., K. Daniilidis, T. Thórhallson, H.-H. Nagel (1992). Model-based object tracking in traffic scenes *in* G. Sandini (Ed.) Computer Vision - ECCV'92. Springer, Berlin/Heidelberg.

Koller, D., K. Danilidis, H. H. Nagel (1993). Model-based object tracking in monocular image sequences of road traffic scenes. International Journal of Computer Vision 10 (3), pp. 257-281.

Larsson, H. (2009). Förarstöd för lägre bränsleförbrukning och minskade emissioner: utvärdering av två system (*in Swedish*) Driver support for better fuel consumption and lower emissions: evaluation of two systems. Licenciate thesis, Lund University, Department for Technology and Society, Faculty of Engineering LTH.

Laureshyn, A., H. Ardö (2006). Automated video analysis as a tool for analysing road user behaviour. ITS World Congress, London, UK.

Laureshyn, A., H. Ardö, T. Jonsson, Å. Svensson (2009). Application of automated video analysis for behavioural studies: concept and experience. IET Intelligent Transport Systems 3 (3), pp. 345-357.

Leibe, B., K. Schindler, N. Cornelis, L. V. Gool (2008). Coupled object detection and tracking from static cameras and moving vehicles. Pattern Analysis and Machine Intelligence, IEEE Transactions 30 (10), pp. 1683-1698.

Maycock, G., P. J. Brocklebank, R. D. Hall (1998). Road layout design and driver behaviour. Transport Research Laboratory (TRL). Report No. 332.

Messelodi, S., C. M. Modena (2005). A computer vision system for traffic accident risk measurement: a case study. Instituto Trentino di Cultura (ITC). Technical Report T05-06-07.

Messelodi, S., C. M. Modena, M. Zanin (2004). A computer vision system for the detection of vehicles at urbam road intersections. Instituto Trentino di Cultura (ITC). Technical Report T04-02-07.

Michon, J. A. (1985). A critical view of driver behaviour models: what do we know what should we know? *in* L. Evand and R. C. Schwing (Eds.). Human behaviour and traffic safety. Plenum Press, New York.

Minderhoud, M. M., P. H. L. Bovy (2001). Extended time-to-collision measures for road traffic safety assessment. Accident Analysis and Prevention 33, pp. 89-97.

Moore, G. (1965). Cramming more components onto integrated circuits. Electronics 38 (8).

Moosmann, F., E. Nowak, F. Jurie (2008). Randomized Clustering Forests for Image Classification. Pattern Analysis and Machine Intelligence, IEEE Transactions 30 (9), pp. 1632-1646.

Nilsson, G. (2004). Traffic Safety Dimensions and the Power Model to Describe the Effect of Speed on Safety. Doctoral thesis, Lund University, Department of Technology and Society, Traffic Engineering.

Nygård, M. (1999). A method for analysing traffic safety with help of speed profiles. Master thesis, Tampere University of Technology, Department of Civil Engineering.

Odelid, K., T. Radjesh, Å. Svensson (1991). Automatisk bildbehandling vid trafikanalyser - en utveckling av den svenska konflikt tekniken (*in Swedish*) Automated image processing in traffic analysis - development of the Swedish conflict technique. Lunds Tekniska Högskola, Institution för Trafikteknik.

Odelid, K., Å. Svensson (1993). Bildbehandling och konfliktteknikken (*in Swedish*) Image processing and the conflict technique. Lunds Tekniska Högskola, Institution för Trafikteknik.

Parkhurst, D. (2006). Using digital video analysis to monitor driver behaviour at intersections. Center for Transportation Research and Education, Iowa State University. CTRE Project 05-214.

Peräaho, M., E. Keskinen, M. Hatakka (2003). Driver competence in a hierarchial perspective: implications for driver education. University of Turku, Traffic Research. Report for Swedish Road Administration.

Quimby, A., G. Maycock, C. Palmer, S. Buttress (1999). The factors that influence a driver's choice of speed: a questionnaire study. Transportation Research Laboratory (TRL). Report No. 325.

Radun, I., J. E. Radun (2009). Convicted of fatigued driving: who, why and how? Accident Analysis & Prevention 41, pp. 869-875.

Reinartz, P., T. Krauss, M. Lachaise, E. Schmeer (2006). Large area traffic monitoring with serial images from airborne cameras. ITS World Congress, London, UK.

Ripley, B. D. (1996). Pattern recognition and neural networks. Cambridge University Press.

Risser, R. (1985). Behavior in traffic conflict situations. Accident Analysis & Prevention 17 (2), pp. 179-197.

Rudolph, V. (1996). Das VideoVerkehrsAnalyse-System VIVAtraffic (*in German*) Video analysis system VIVAtraffic *in* H. H. Topp (Ed.) Traffic safety work with video processing. Transportation Department, University Kaiserslautern.

Räsänen, M. (2005). Effects of a rumble strip barrier line on lane keeping in a curve. Accident Analysis & Prevention 37 (3), pp. 575-581.

Saunier, N., T. Sayed (2007). Automated analysis of road safety with video data. Transportation Research Record: Journal of the Transportation Research Board 2019, pp. 57-64.

Schoenemann, T., D. Cremers (2008). Globally optimal shape-based tracking in real-time. IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, Alaska.

Sekine, T., E. Sekine (2009). Characteristic classification of driving behaviour with mixed traffic situation for driver assist. 21st International Symposium on Dynamics of Vehicles on Roads and Tracks, Stockholm.

Shbeeb, L. (2000). Development of a traffic conflicts technique for different environments - a comparative study of pedestrian conflicts in Sweden and Jordan. Doctoral thesis, University of Lund, Lund Institute of Technology, Department of Technology and Society, Traffic Engineering.

Shi, J., C. Tomasi (1994). Good features to track. Computer Vision and Pattern Recognition. Proceedings of IEEE Computer Society Conference on CVPR, pp. 593-600.

SIKA (2009). Road traffic injuries 2008. Statens institut för kommunikationsanalys. SIKA Statistik 2009:23.

Song, X., R. Nevatia (2007). Detection and tracking of moving vehicles in crowded scenes. IEEE Workshop on Motion and Video Computing.

Strang, G. (1986). Introduction to applied mathematics. Wellesley-Cambridge Press, Wellesley, Massachusetts.

Svensson, Å. (1992). Vidareutveckling och validering av den svenska konfliktteknikken (*in Swedish*) Further development and validation of the Swedish Traffic Conflicts Technique. Department of Traffic Planning and Engineering, Lund University.

Svensson, Å. (1998). A method for analysing the traffic process in a safety perspective. Doctoral thesis, University of Lund, Lund Institute of Technology, Department of Traffic Planning and Engineering.

Svensson, Å., C. Hydén (2006). Estimating the severity of safety related behaviour. Accident Analysis & Prevention 38, pp. 379-385.

Tan, T. N., G. D. Sallivan, K. D. Baker (1998). Model-based localisation and recognition of road vehicles. International Journal of Computer Vision 27 (1), pp. 525.

Thompson, S. E., S. Parthasarathy (2006). Moore's law: the future of Si microelectronics Materials Today 9 (6), pp. 20-25.
Traficon® (2008). Video for traffic data acquisition, automatic incident and presence detection. www.traficon.com *Acc.* 9 Feb 2009.
Transportstyrelsen (2009). STRADA – informationssystem för olyckor & skador (*in Swedish*) STRADA - infornation system for accidents and injuries. http://www.transportstyrelsen.se/sv/Vag/STRADA-informationssystem-for-olyckor-skador/ *Acc.* 14 Sep 2009.
van der Horst, R. (1990). A time-based analysis of road user behaviour in normal and critical encounters. Doctoral thesis, Delft University of Technology.
van der Horst, R., R. J. Kraay (1986). The Dutch conflict technique – DOCTOR. ICTCT Workshop, Budapest.
van Winsum, W., K. A. Brookhuis, D. de Waard (2000). A comparison of different ways to approximate time-to-lane crossing (TLC) during car driving. Accident Analysis and Prevention 32, pp. 47-56.
Várhelyi, A. (1998). Driver's speed behaviour at a zebra crossing: a case study. Accident Analysis and Prevention 30 (6), pp. 731-743.
Várhelyi, A., M. Hjälmdahl, C. Hydén, M. Draskóczy (2004). Effects of an active accelerator pedal on driver behaviour and traffic safety after long-term use in urban areas. Accident Analysis and Prevention 36, pp. 729-737.
Wender, S., K. Dietmayer (2007). 3D vehicle detection using laser scanner and a video camera. 6th European Congress on ITS, Ålborg, Denmark.
Wilde, G. J. S. (1994). Target Risk. PDE Publications, Toronto.
VisioWay (2007). Video detection systems. www.visioway.com *Acc.* 9 Feb 2009.
Vogel, K. (2002). What characterizes a "free vehicle" in an urban area? Transportation Research Part F 5, pp. 15-29.
Vägverket (2008). Målstyrning av trafiksäkerhetsarbetet (*in Swedish*) Management by objectives for road safety work. Swedish Road Administration. Publikation 2008:31.
Xu, F., X. Lia, K. Fujimura (2005). Pedestrian detection and tracking with night vision. IEEE transactions on intellegent transportation systems 6 (1), pp. 63-71.
Yung, N. H. C., A. H. S. Lai (2001). An effective video analysis method for detection red light runners. IEEE transactions on vehicular technology 50 (40), pp. 1074-1084.
Zhang, L., V. Kovvali, N. Clark, D. Sallman, V. Alexiadis (2007). NGSIM-VIDEO User's Manual. Department of Transportation, Federal Highway Administration. FHWA-HOP-07-009.
Zhang, L., Y. Li, R. Nevatia (2008). Global data association for multi-object tracking using network flows. IEEE Conference on Computer Vision and Pattern Recognition.
Zhao, T., R. Nevatia (2003). Car detection in low resolution aerial images. Image and Vision Computing 21, pp. 693-703.
Zhao, T., R. Nevatia (2004). Tracking multiple humans in complex situations. Pattern Analysis and Machine Intelligence, IEEE Transactions 26 (9), pp. 1208-1221.
Åström, K., J. Karlsson, O. Enquist, A. Ericsson, F. Kahl (2007). Automatic Feature Point Correspondences and Shape Analysis with Missing Data and Outliers *in* B. K. Ersböll and K. S. Stenstrup (Eds.). Lecture Notes in Computer Science. Springer.

Paper I

Submitted to "Journal of ITS" on 25th June 2008
Re-submitted after a revision on 11th September 2009

ROAD USER BEHAVIOUR INDICATORS IN AUTOMATED VIDEO ANALYSIS SYSTEMS

Aliaksei Laureshyn, Åse Svensson
Traffic and Roads, Department of Technology and Society,
Faculty of Engineering LTH, Lund University
Box 118, 22100 Lund, Sweden Tel. +46 46 222 91 25 Fax +46 46 12 32 72

Abstract: Many automated video analysis systems are available today, but very few of them have the direct functionality necessary to make them usable in studies of behaviour in road traffic. This paper provides a perspective on the factors that affect the applicability of video analysis in this area, and attempts to promote a better dialog between the developers of video analysis systems and traffic behaviour researchers. A system developed at Lund University, Sweden, is taken as a typical example of video analysis research with a traffic application, and its capabilities, performance quality, and advantages and limitations are illustrated and discussed. The paper also discusses the nature of an indicator and the requirements, from validity and reliability perspectives, for "good" indicators in behavioural research. It presents an overview of some recent traffic research reports, provides a list of indicators currently used and estimates their applicability in automated video analysis. The main conclusion is that most of the reviewed indicators can be retrieved from video data. In current practice the majority of indicators are of "binary" (yes/no) or "single value" types. Video analysis makes it possible to collect "continuous" indicators to a larger extent, which enables more sophisticated analysis of the data and emphasises the prerequisites for the chosen indicators to be valid representatives of the qualities of interest (e.g. safety or efficiency).

Keywords: Automated video analysis, indicators, reliability, road user behaviour, validity

1. INTRODUCTION

Road user behaviour, which is an important research area in traffic science, helps to explain "why road users act as they do and explores possibilities for influencing their actions" (*Englund et al., 1998*). Behavioural studies differ from other attempts to apply psychological theories, since factors like "personality", "attitude", "motivation", etc. are considered too complex and the available knowledge about them too limited to be able to measure them in a reliable way; therefore, the efforts are concentrated on studying the behaviour itself. When applied in traffic research, this means exploring how individual road users act in certain traffic conditions (described by measurable factors like road design, frequency and types of interactions with other road users,

etc.) and how their actions change with changes in the conditions. This provides practical knowledge of the factors that promote the desired behaviour patterns, which is vital for the planning, design and successful management of a traffic system.

If behaviour is to be measured in a practical and applicable way, its description has to be restricted to some indicators. Thus, a decision on what are relevant indicators is very important. The ultimate purpose of behavioural research in this context is to be able to relate certain types of behaviour to the qualities of the traffic environment, i.e., safety, efficiency and comfort for road users. Therefore, a theory explaining why a behavioural indicator reflects those qualities must be found and its validity has to be tested. It should be possible to measure the indicator with available techniques, and their definition should preferably be in line with the indicators used in other similar studies. So far behavioural studies in traffic are often conducted on a small scale and restricted to few locations and short observation periods. Obtaining a better understanding of the studied phenomena often necessitates comparing the results or even merging data from different studies, which is only possible if indicators used in each study are defined and applied in the same way.

Conducting a behavioural study (for example, using a conflict technique) as an evaluation after the re-design of a road infrastructure is an effective way to measure the extent to which the expected changes take place, as well as detect bad functioning and any side-effects in the system at an early stage. Consequently, it should be promoted as a standard procedure. Still, evaluation studies are still far from being common practice (*Elvik, 1997*), the main reasons being inefficiency and high costs of data collection, which in most cases is done manually by human observers.

Automated video analysis is a tool capable of collecting data on road user behaviour in a more efficient and systematic way. The use of video recordings has many advantages, such as the possibility of making long-period observations (which is crucial if the frequency of the events is low), an unobtrusive way of collecting data and the opportunity to look through the relevant situations again or study very complex situations in detail. Automation of the video analysis allows us to avoid watching uninteresting videos and concentrate on the relevant situations.

Nonetheless, it appears that communication between the developers of the video analysis systems and traffic behaviour researchers is still not properly established. Most often the tasks performed by video analysis systems are limited to detection of road users and extracting their trajectories, but the scope of behavioural indicators is much wider than that (e.g. *Fernandez-Caballero et al., 2007, Suzuki & Nakamura, 2006, Messelodi et al., 2004*). Even if more advanced indicators are tested, they are not always related to the knowledge already existing and obtained from previous behavioural studies and thus their validity is questionable (e. g. *Saunier & Sayed, 2007, Atev et al., 2005, Messelodi & Modena, 2005*). There is an enormous focus on the detection of motor-vehicles, while pedestrians and cyclists, who play as important a role and are as frequent as vehicles, at least in urban environments, are often ignored. On the other hand, behavioural studies employing the advantages of video analysis techniques are rare despite the fact that many systems are now available and

tests show that their performance is quite satisfactory (*Laureshyn et al., 2009, Malkhamah et al., 2005*).

This paper is an attempt to promote a dialog between the developers of video analysis systems and traffic behaviour researchers. It illustrates the capabilities and limitations of automated video analysis, and focuses on the qualities that are required for successful use of this technique in behavioural studies. It contains an overview of the indicators currently used in traffic behaviour research, and takes up the question of whether relevant data can be collected using video analysis systems.

2. AUTOMATED VIDEO ANALYSIS SYSTEMS

Video recording is a method commonly used for data collection in traffic behaviour research, either on its own or as a complement to other measurement methods (*Archer, 2005, Malkhamah et al., 2005, Räsänen, 2005, Andersson, 2000, Svensson, 1998, van der Horst, 1990*). Still, video recording itself does not change the difficulties associated with human observations as a detection method. It is still the observer, either located at a roadside or sitting in front of a monitor, who makes the necessary measurements and detects the events of interest. Watching a video film takes at least the same amount of time as on-site observations, while the results in some cases might be somewhat inaccurate since video films cannot reflect the traffic environment completely.

Automated video analysis utilises computer techniques that can detect and track moving objects caught by the camera. There are numerous systems developed with a focus on car drivers' surveillance, e.g. *Ji et al., 2006, Parkhurst, 2006, Sakamoto et al., 2006, Messelodi et al., 2004, Coifman et al., 1998*, and some are available as commercial products (*Autoscope®, 2009, Hitachi, 2008, Traficon®, 2008, VisioWay, 2007*). These systems are often applied in less complex situations such as on motorways with most vehicles driving straight ahead, no crossing manoeuvres or mixture of different types of road users (e.g. no pedestrians and cyclists are present or they are not detected) and the performed tasks are also quite simple – detection of congestions, traffic counting, etc. There are also some solutions available that can detect and track pedestrians in environments with only pedestrians present, such as parks or walkways (*Berclaz et al., 2006, Zhao & Nevatia, 2004, Isard & MacCormick, 2001*, commercially available pedestrian counters *Cognimatics, 2009, Springboard, 2009*). In some cases ignorance of road user type or limitation of the distinguished types to very few allows the analysis to be performed in real time and the equipment to be quite compact and relatively inexpensive, which explains the fast spreading and growing popularity of such systems.

In environments where road users of different types are mixed, it is necessary to make a decision on what type of road user is to be detected. Most often it is done by manually specifying the size parameters for wire frame models or by training the system on a large number of manually classified examples (*Leibe et al., 2008, Schoenemann & Cremers, 2008, Tan et al., 1998, Koller et al., 1992*). Other methods work with coarser models where it is enough to specify some approximate size of the road users (*Song & Nevatia, 2007*).

A research group at the Faculty of Engineering, LTH, Lund University, is developing a system specially aimed at studying the behaviour of road users in complex traffic environments (primarily urban conditions and mixed traffic modes). The video analysis part of the system is a relatively typical example of research in this area and we will use it as an example to investigate the applicability of such systems in behavioural studies. The technical description is intentionally kept short and for further details the reader is referred to *Ardö, 2009* and *Laureshyn et al., 2009*.

The system uses video recorded from above the scene (typically, from 12-25 meters height) and includes several techniques, which can be chosen depending on what types of events are to be detected and what accuracy is required. The initial detection of moving objects is done by using a KLT (Kanade-Lucas-Tomasi) interest point tracker (*Shi & Tomasi, 1994*). The tracker produces trajectories of the easily distinguishable points, such as corners or edge junctions (typically, several points belong to one road user), which can be used as a simple detector of movements in a pre-defined direction, for example, driving against a prescribed direction along a one-way street. If it is known which tracks belong to a road user, the set of points in each frame can be used to estimate the road-user position (the mean value of the points is most often a point close to the centre of the road user). Comparing the tracks with the results of background/foreground segmentation (i.e. selection of the areas in each frame that show the moving objects, the foreground, and not the static road environment, the background, *Ardö & Åström, 2008*) allows this decision to be made. The total size of the foreground components representing a single road user in pixels is correlated to the road user's size in the real-world (a bus most often covers a larger area in the image than a pedestrian) and can be used for rough type classification.

The direct speed calculations at a differentiated position yield quite inaccurate results since the position, estimated as a mean of the interest points, jumps a bit back and forward as new points are found and the old ones are lost. Better precision is obtained if a road user is considered as a dynamic set of interest points, which is subject to rotation, translation and scaling (*Åström et al., 2007*). This model no longer assumes that the same interest points are available all the time and does not create jumps when the points are lost or added. It does not take into account the fact that the interest points come from 3-dimensional objects projected on 2-dimensional images, therefore the estimated position might quite often appear outside the road user borders. However, differentiation of this position provides much smoother and accurate speed estimates.

A major drawback of one-camera techniques is that it is not easy to restore the 3-dimensional shape of a road user from the 2-dimensional information of one camera without making approximating assumptions. If a camera is located high enough over the road, a reasonable assumption is that road users are flat and belong to the road plane. The downside is that this introduces a systematic error in position estimation depending on the road user's height, orientation and distance from the camera, i.e., it is not constant even for the same road user during a passage. It does not, however, change significantly between two adjacent frames and therefore the estimates of speed are not affected (*Laureshyn et al., 2009*).

If video recordings are available from more than one camera, the 3-dimensional information can theoretically be restored. If the cameras are synchronised (frames are taken strictly at the same time), the differences in position of the same point in the images from different cameras allow calculation of the "depth" of the point in the image (stereo vision, *Laureshyn & Ardö, 2006*). However, the requirements for cameras installation (2 mounting points with 10-15 meters in between, with a good view and available power supply, synchronising cable between the cameras, etc.) appeared to be too impractical to be widely used and it was not pursued in these tests.

Another solution is a model where road users are approximated with a set of "boxes" with pre-defined dimensions, and the entire analysed scene is described by its state, consisting of the location, orientation and types of all road users present at each moment of time. Then, a set of state hypotheses is generated. Each of these hypotheses can be matched to the observations made by a camera(s) surveying the scene. The states can be combined into sequences by designing a Hidden Markov Model (HMM), which can be optimised over state sequences. Thereby the state sequence best describing the observations made by all the cameras is found, and it is this state sequence that contains the trajectories of all the road users. The advantage of this technique is that it has fewer requirements for camera installation (it is sufficient that a camera covers just a part of the studied area as long as the cameras overlap; no need for exact time synchronisation of the cameras). However, as the number of possible states (depending on the number of considered road user types, cell size of the positions grid, etc.) is increasing, this technique is becoming more and more computationally intensive (*Ardö, 2009*).

From a user perspective, the most important questions about a video analysis system are: i) how easy is it to collect the video data? ii) how long does it take to process the data? iii) what is the accuracy of the results in terms of detection rate, position and speed estimates? The system has been used in two large-scale studies and the experience gained provides some answers to these questions. Study I (*Laureshyn et al., 2009*) deals with safety effects of cycling against traffic along one-way streets and is based on video material recorded at 22 intersections, 3-4 days each. Study II (*Sakshaug et al., 2009*) compares two roundabouts that differ in the way their designs treat the cyclist flows; 5 days of video recordings are analysed for each site.

2.1. Data collection

The installation of a camera requires several practical problems to be solved. The camera location must provide a good view, with a proper angle, of the scene. Longer recordings cannot be run on batteries and therefore access to a stationary power supply is necessary. It is an advantage for the camera to be installed on a piece of infrastructure already existing (e.g. balcony railings of a nearby house or a lamp post) so that a new large object like a mobile mast does not distract road users a lot. Even slight swaying of a mast because of wind or road vibrations makes the video "shake" and has to be compensated for before the video analysis algorithms can be run, which increases the processing time.

Finding a good place for cameras turned out to be a big problem and took a very long time for both studies. In Study I, 32 sites were planned for filming, but 10 of them were cancelled as no suitable positions for the cameras were found. Four more sites were skipped later as the cameras turned out to be too far away from the scene and proper analysis was not possible. In Study II several flows coming from different directions were of interest, but it was not possible to get the best view of all of them from one point. Finally, it was decided to set the camera so that one of the directions got the best view, and to sacrifice the others, which resulted in clear differences in the detection quality of the directions.

The quality of the recordings was also affected by factors like unfavourable weather conditions (heavy rain, snow, fog) and sun glare in the morning and evening hours. Recordings done in the dark differs quite a lot from the daylight conditions and required special adjustments of techniques for analysis (e.g. the light patch on asphalt in front of a car with headlamps on is often detected as a separate object).

2.2. Processing time

Generally, the more advanced a technique is, the more manual settings and longer computation time it requires and the more sensitive it is for possible input errors. Table 1 shows the relation between the length of input video and time necessary to process it for several of the techniques described earlier. The computation time is highly dependent on what hardware is used; therefore the figures should be considered as relative and used mostly to compare the techniques.

Table 1. Computation time for different video analysis techniques.

Video analysis technique	*Output data*	*No cameras*	*Computation time (2.40 GHz, Pentium IV)*
interest points tracker	- detected objects moving in a pre-defined direction	*1*	*2* hours per *1* hour input video
point tracker combined with foreground/background segmentation	- trajectories and speed profiles; - road user types (based on size criteria)	*1*	*18* hours per *1* hour input video
HMM state model		*1+*	*96* hours per *1* minute input video (small test, *4* cameras used)

Processing recordings of longer periods takes a significant amount of time. However, once the initial settings are done, the calculations run automatically and do not require an operator. There are several possible ways to decrease the waiting time for the results. Splitting the computational tasks among several computers is the most straightforward and effective way. The calculations themselves can be speeded up by updating the hardware and rewriting parts of the code in a low-level programming language.

2.3. Accuracy

The quality of the detections depends on the algorithms used, but also on many other factors like the camera view, resolution and quality of the input video, lighting and weather conditions, traffic density and, as these conditions are not constant over time, the detection rate also varies.

Table 2 shows the results of the detection quality tests done in Studies I and II. One-camera recordings and interest point-based trackers of road users were used in both studies and the goal was to detect cyclists moving in certain directions. To check the detection quality some parts of the videos were looked through and cyclists were detected manually. This was expected to provide the "ground truth". In Study I, however, the video analysis found more cyclists than a human observer first did at some sites with very intensive and mixed pedestrian and cyclist flows. In those cases the automated detection numbers were taken as the "ground truth". There are some concerns that the "ground truth" in the table does not represent the actual number of cyclists as some of them might still have been missed by both the observers and the system.

Table 2. Detection tests, Study I and II (adopted from *Laureshyn et al., 2009* and *Sakshaug et al., 2009)***.**

Study, technique used	***Observation time***	***Location***	***Cyclists, "ground truth"***	***Automated detection from video***				
				Detections, total	***Cyclists***	***Detection rate****	***False positives***	***False positive rate*****
Study I, point tracker (one camera)	0.5-1 hour each site, 15 hours totally	18 sites, total	*86*	*268*	*60*	*70%*	*208*	*78%*
Study II, trajectory extraction (one camera)	9 hours	Site 1, direction 1	*149*	*150*	*98*	*66%*	*52*	*35%*
		Site 1, direction 2	*242*	*152*	*104*	*43%*	*48*	*32%*
		Site 1, direction 3	*99*	*51*	*31*	*31%*	*20*	*39%*
		Site 1, direction 4	*177*	*40*	*32*	*18%*	*8*	*20%*
	9 hours	Site 2, direction 1	*541*	*585*	*387*	*72%*	*198*	*34%*
		Site 2, direction 2	*172*	*791*	*82*	*48%*	*708*	*90%*
		Site 2, direction 3	*832*	*1378*	*485*	*58%*	*895*	*65%*

* Detection rate is the ratio between the number of correct detections and the "ground truth".
** False positive rate is the ration between the number of false positives and the total number of detections.

In Study II it was not possible to get the optimal view for all the locations, and therefore the camera orientation was optimised to view the directions 1 and 2 at Site 1 and direction 1 at Site 2. This partly explains the poor detection quality in other directions.

A special test was performed to check the accuracy of the trajectory and the speed estimates. A vehicle (passenger car) equipped with a speed logger and a high-precision GNSS receiver[1] drove through an intersection where four cameras were installed and filmed from these four locations and angles. Figure 1*a* compares the vehicle's trajectory estimated from the interest points using one-camera data (green) and from the HMM state model employing several cameras' data (red) with the GNSS measurements. Plotted on a single frame image, it is clearly seen how the estimate from the interest point tracker is shifted from the real position (Figure 1*b*) as compared to the estimate based on the HMM state model (Figure 1*c*).

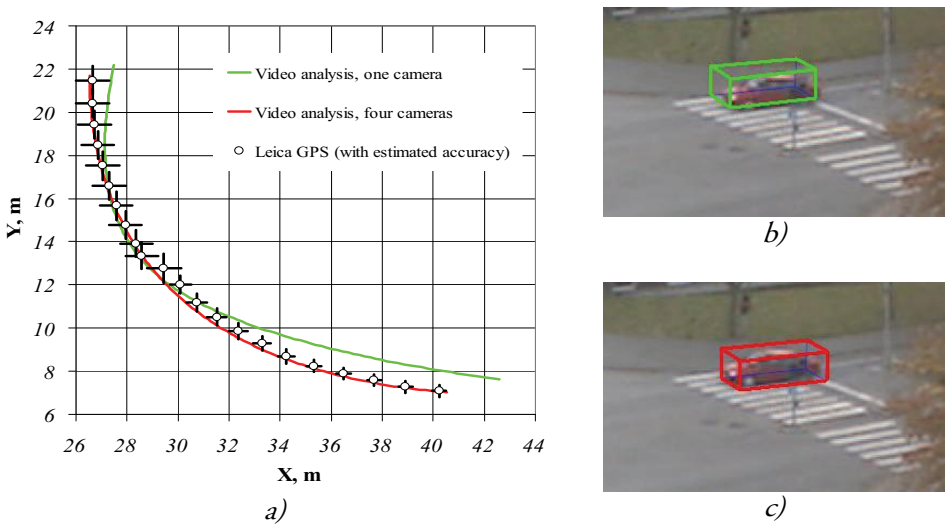


Figure 1. Accuracy tests of position estimates by two video analysis techniques: black circles – from Leica GNSS (with shown estimated accuracy of measurements taken, rms), green – based on data from one camera and interest point tracking algorithm, red – based on data from several cameras and HMM state model. a) trajectory; b), c) position estimates at a single frame.

Figure 2 compares the speed profile estimated from one camera data with the speed log from the vehicle. The procedure for speed estimation using the HMM state model has not yet been properly developed and tested.

[1] Leica GX1230 GG, http://www.leica-geosystems.com/. In ideal conditions the provided accuracy is 10-20 mm, but factors like occluded satellites, multipath, ionospheric conditions, etc. may decrease the accuracy significantly. The actual estimated accuracy varied from 5 to 70 cm (rms) in the test.

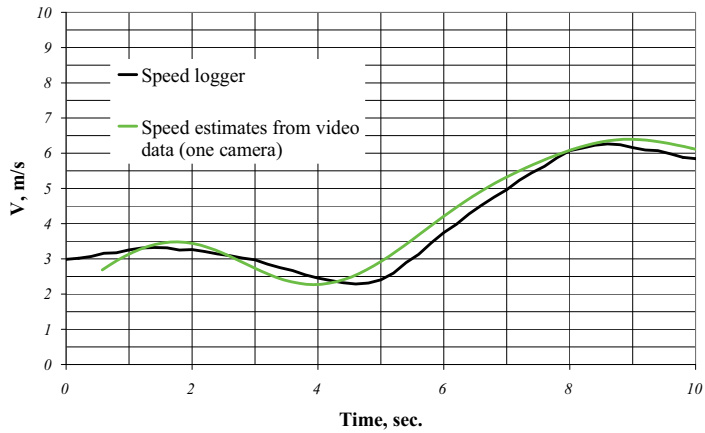


Figure 2. Speed estimates from the video data (one camera).

3. INDICATORS IN A BEHAVIOURAL STUDY

For a start, let us construct a schematic chain to describe the interactions between a road user and the traffic environment (see Figure 3). A road user (car driver, pedestrian, cyclist, etc.) is affected by many factors: (i) those related to the road environment (e.g. visibility, road surface quality, other road users, etc.), (ii) legislative and social norms regulating the activity in this environment and (iii) personal factors (such as knowledge and experience, physiological ability to react adequately and in good time, fatigue, etc.). All these factors together form a particular behaviour (actions) of the road user, which in turn causes some changes in the situation – the outcome. The new situation also becomes an influential factor for the road user, thus completing the chain.

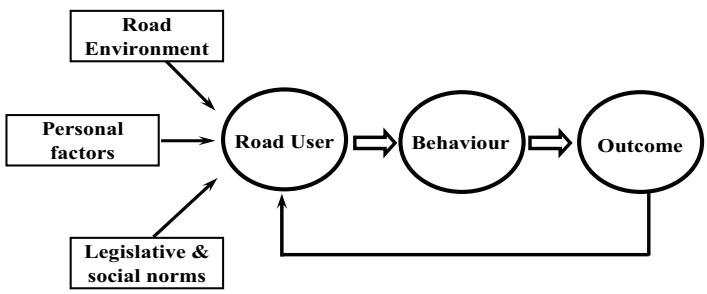


Figure 3. Illustration of the interaction process.

Road user behaviour is a many-sided and complicated phenomenon and the list of indicators describing it can theoretically be endless. However, measuring "everything", apart from being very resource-consuming, does not necessarily guarantee the credibility of the results since the indicators have to be further interpreted in terms of certain qualities of the traffic system (e.g. safety). The two key issues when deciding if an indicator is a good representative of the desired quality, and thus should be measured, are its validity and reliability:

(i) The ***validity*** refers to whether an indicator describes the quality that it is intended to represent and to what extent. For example, the high number of red-walkers at a pedestrian crossing may be interpreted as lack of respect for the traffic rules, ignorance or unawareness of the risks, thus indicating the "unsafety". An alternative interpretation might be the general feeling of security in a traffic system that allows small violations without severe consequences, i.e., it is the indicator of "safety". Establishing the validity of an indicator usually requires numerous large-scale studies performed in various conditions. The correlations found have to be supported by a theory providing clear logical and causal connection between the indicator and the quality it is supposed to represent.

(ii) ***Reliability*** refers to the methods used to measure the indicator and the accuracy of the measurement. Ideally, a reliable method produces the same results irrespective of who performs the measurement, and the accuracy remains the same for any measurement location and any conditions, thus ensuring that the difference in the results reflects the difference in the studied phenomenon and not just a fluctuation in the measurements accuracy. For example, traffic conflict techniques, which were much criticised for being unreliable as conflict detection completely relies on a human observer's subjective judgements of road users' speeds and distances between them. Later, tests comparing the estimations of different observers showed that the produced results were very similar, i.e., the method is quite reliable (*Hydén, 1987*).

An indicator may describe different parts in the illustrated interaction chain (Figure 3). Based on this criterion, the following classification is henceforth used in this paper:

I. *Road user indicators* – describing some properties of an individual road user (e.g. age, gender, etc.);
II. *Individual behaviour indicators* – describing the behaviour of a single road user;
III. *Interaction indicators* – describing the process of interaction of several road users;
IV. *System functioning indicators* – describing the traffic situation in general (the outcome).

From society's perspective the main interest lies in the outcome of road users' behaviour, i.e., we want to know how safe or effective the whole system is. However, in many cases the outcome is not that obvious at a first glance or it might take a significant amount of time to get a representative estimate of the outcome indicators. Thus, it has been stressed (e.g. *Svensson, 1998*) that there is a need to use some other methods to predict the outcome, for instance, using the indicators of the preceding links in the chain – the behaviour and interaction indicators. The transfer from one link to the other implies further uncertainty and thus puts even higher demands on the validation of the indicators selected for use.

4. REVIEW OF LITERATURE ON BEHAVIOUR INDICATORS

The following sections provide a brief overview of some research studies on individual road user behaviour, interactions and a specific type of the latter, traffic conflicts. The studies are grouped with regard to the main road user group on which the study is

focused (drivers of motor vehicles, cyclists or pedestrians). The purpose is to create a snapshot list of indicators currently used in traffic research – the actual findings and conclusions of the studies are thus not discussed at all.

4.1. Study design

The main criteria for a study to be selected are that it is relatively recent and contributes new indicators to those already on the list; therefore there are many other reports related to specified behaviour that are not included here. Some indicators are quite universal (e.g. Time-to-Collision) and may be applied when studying any type of road users. To avoid redundancy, they are mentioned just once and comments are made on their universality.

In the summary tables (Table 3-6) "type of indicator" is assigned according to the classification above (I to IV). The "data type" is expressed as either "*binary*" ("yes/no", e.g. if a car stops or not), "*single value*" (e.g. pedestrian age) or "*sequential*" (a set of values, e.g. speed profile over time).

The position of a road user, expressed as X-Y co-ordinates, as well as some other parameters like speed, size and orientation are the most typical output from the video analysis systems (Figure 1). Therefore the applicability of an indicator in video analysis-based studies is highly dependent on whether the indicator can be expressed through these parameters or not. The classification of applicability is done by assigning one of the four types: "*X-Y*", which means that the indicator can be calculated from the co-ordinates and related parameters only; "*X-Y, compl.*" (requires complementary input, for example information from the traffic light on what signal is currently active); "visual control" (cannot be calculated from the co-ordinates, but if the situation is detected by other criteria an observer can retrieve the indicator when watching the video clip); and "*poor*", which means that retrieving the indicator from the video is problematic (e.g. the presence of eye contact between drivers).

4.2. Driver behaviour and driver-driver interactions

The interaction between two drivers (or road users of any other type) is often described by a *Time-to-Collision* (TTC) parameter, which is defined as the time required for two vehicles to collide if they remain at their present speed and on the same path (*Hayward, 1971*, cited at *van der Horst, 1990*). This definition points out the requirement for the vehicles to be on a *collision course* in order to calculate TTC, which becomes infinite if the vehicles' trajectories do not cross or they pass the common spatial zone with a time difference.

TTC is a continuous parameter and may be calculated for any moment as long as the collision course requirement is fulfilled. In practice, though, an encounter is often characterised by only one value, for example, the minimal TTC_{min} or whether TTC reaches a set critical value TTC^{*} or not. *Minderhoud & Bovy, 2001* propose two extended TTC measures. The *Time Exposed Time-to-Collision* (TET) is defined as the "duration of exposition to safety-critical Time-to-Collision values over specified time duration" and may be used to assess an individual interaction or as an integrated parameter, describing all the interactions of a vehicle or all vehicles on a road section

within a specified time. *Time Integrated Time-to-Collision* (TIT) is intended to take into account not only the duration, but also the magnitude of the TTC going below the critical value, and is calculated as an integral of the TTC-profile below the TTC^*:

$$TIT = \int [TTC^* - TTC_i(t)]dt \ , \ 0 \le TTC_i(t) \le TTC^* .$$

van der Horst, 1990, discusses whether the TTC calculation can be enhanced by assuming constant longitudinal acceleration and angular velocity, which represent the situation when the driver does not react anymore and the steering wheel and the accelerator remain at fixed positions. While the assumption of constant acceleration appears to be quite reasonable (and TTC calculated in this manner is referred as *TTCA*), this is not the case for the constant angular velocity assumption where the estimated trajectories easily take unrealistic shapes and go off the road. It is also proposed that TTC may be calculated for a particular moment (e.g. start of braking, TTC_{br} or driver's first head movement, TTC_{1st}) or, instead of the collision point, the time may be estimated to the moment the vehicle reaches some other point or object (e.g. intersection stop line, referred to as *TTI*).

In a search for effective indicators of drivers' last-second evasive braking in rear-end encounters, *Kiefer et al., 2005*, propose an *Enhanced TTC* parameter (ETTC), which takes into account the deceleration of the lead vehicle. Among other tested indicators the *inverse TTC* (1/TTC) is the best predictor to distinguish between normal and hard last-second braking.

Post-Encroachment-Time (PET) is a measure which represents the time difference between two vehicles passing the common spatial zone and, thus, becomes zero when the vehicles are on a collision course. Unlike TTC, PET can be measured directly, which makes it potentially a more convenient indicator for practical use. However, if the common zone is not clearly defined (as in a car-following situation), the calculation of PET is impossible (*Allen et al., 1977*).

Gettman & Head, 2003, propose some further derivatives of PET (applied to the encounter between left-turning and driving-through vehicles):

"Gap Time – the time lapse between the completion of encroachment by a turning vehicle and the arrival time of a crossing vehicle if they continue with the same speed and path.

Encroachment Time – time duration during which the turning vehicle infringes upon the right-of-way of the through-driving vehicle.

Initially Attempted Post-Encroachment Time – time lapse between commencement of encroachment by the turning vehicle, plus the expected time for the through-vehicle to reach the point of collision, and the completion time of encroachment by the turning vehicle.

van Winsum et al., 2000, van Winsum et al., 1999, applied the TTC concept to study the lane-following and lane-changing manoeuvres and introduced the *Time-to-Line Crossing* (TLC) indicator. TLC is calculated as the time available for the driver until the moment at which any part of the vehicle reaches one of the lane boundaries.

Low TLC values are proposed for use as triggers, activating the lane-following assistance system.

Another parameter, describing the interaction between two vehicles in a car-following situation is *time headway*, defined as "the elapsed time between the front of the lead vehicle passing a point on the roadway and the front of the following vehicle passing the same point" (*Evans, 1991*). *Vogel, 2003*, compares the use of time headway and TTC and concludes that these parameters do not substitute for each other, but provide different information and may be used for different purposes.

The concept of "*free vehicle*", i.e., a vehicle driving at a desired speed and not affected by other road users, is widely used in traffic research. *Vogel, 2002*, claims that free vehicles may be detected based on the size of the time headway and proposes a headway threshold value of 6 seconds.

The *time gap* is calculated as the headway, but the time is estimated between the back of the leading vehicle and the front of the following one passing the same point (*Vogel, 2002*). Time gap and *gap acceptance* are the key concepts used in intersection capacity calculation and estimation of the Level-of-Service (*Hagring, 2000, TRB, 2000*).

Summala, 1980, studied the effects of the overtaking manoeuvre on the safety margins kept by drivers to other vehicles and road environment elements. The headway to the lead vehicle and the *lateral position,* measured as a distance from the left wheel to the road's middle line, were registered.

Driver behaviour at non-signalised intersections is often described by *compliance with the yielding rules* and *STOP sign requirements* (*Björklund & Åberg, 2005*). At signalised intersections the number of *red runners* is considered as an important safety indicator (*Martinez & Porter, 2006*). In the interaction between a driver and a traffic light the *dilemma zone* is often discussed, i.e., the zone where the driver must make a decision on whether to proceed through the intersection or to stop when the signal changes from green to amber (*Moon & Coleman III, 2003*). More discussion on this can be found in *Köll et al., 2004*.

In a study on how driver performance is affected by fatigue (*Dingus et al., 2006*) triggering limits were set on *steering* (the angular speed of steering wheel turning), *lateral* and *longitudinal accelerations*, Time-to-Collision; other triggers included *lane deviation* (driver crosses the lane border) and *lane departure directly corrected by steering.* Several video cameras constantly observed the driver and road environment, and if any of the trigger parameters exceeded the set limit value, a short video cut before and after the trigger activation was saved. The drivers' *gender* and *age* were taken into consideration when choosing the subjects for the experiment.

Hancock et al., 2003, studied the negative effects of telephone use while driving and assessed both the quality of driving performance and operating the telephone. The driver's task was to drive a loop and stop at a signalised intersection, while he/she was distracted by having to perform some additional tasks with the telephone. The quality of the driving through the intersection was assessed by the *stopping distance* (SD – distance between the stop line and the vehicle front after it has come to a complete halt), *brake response time* (BRT – time between the signal switch and start of the

braking), *stopping time* (ST – time between the first activation of the brake pedal and the moment when the vehicle reached zero velocity), *stopping accuracy* (SA – if the driver successfully completed the stop task or not), *distracter response time* (DRT), *distracter response accuracy*. Drivers were also instructed to maintain a certain *intersection approach speed*.

Table 3 summarizes the indicators mentioned above and shows them as a register.

4.3. Pedestrian behaviour and driver-pedestrian interactions

Interactions between vehicle drivers and pedestrians occur when a pedestrian crosses or aims to cross the road when there is a vehicle approaching. The most common indicators used to describe pedestrian behaviour at signalized crossings are *red light violations* and *pedestrian waiting time* (e.g. *Harré & Wrapson, 2004*). Studying the behaviour of adult pedestrians accompanied by a child, *Zeedyk & Kelly, 2003*, registered *crossing within the marked crosswalk*, *stopping at the curb*, *pressing the button to operate the light*, *waiting for the green light* before beginning to cross, *checking for approaching traffic* and *walking* (rather than *running* across the street). Due to the specific focus of the study (parental influence over the children), additional parameters were *oral introduction to the child*, *holding the child's hand* and *giving the child a chance to press the button her/himself*.

In a study of the effects of a warning system on a pedestrian crosswalk (flashing lights embedded in the pavement, activated when a pedestrian is detected), *Hakkert et al., 2002*, used the following indicators: a) *vehicle speed*; b) *driver yielding to pedestrian(s)*; c) *conflict in the driver-pedestrian interaction*; d) *pedestrian crossing the road outside the crosswalk area*; e) pedestrian keeping to safe crossing rules. *Vehicle speeds* were measured for different *vehicle types* at the approach (at a distance of about 30 m from the crosswalk) and when reaching the crosswalk. When considering the driver's yielding behaviour, the following pedestrian "statuses" were distinguished: a pedestrian *approaching along a sidewalk*, a pedestrian *just stepping off the sidewalk* and a pedestrian *in the middle of the pedestrian crossing*. The *conflict* in the interaction was defined as an abrupt change of course or speed by either a driver or a pedestrian in order to avoid a collision. Pedestrians crossing the road outside the crosswalk were counted within a 30 m zone before and after the marked crosswalk area. Safe pedestrian behaviour implied *stopping* and *looking out for oncoming traffic* before stepping onto the road, where the following traffic situations were distinguished: *no oncoming traffic* and *vehicle approaching on a lane close to/far from the pedestrian* (the study was focused on one half of a crosswalk, i.e. the case of one-way vehicle traffic and two-way pedestrian traffic).

Table 3. Indicators describing drivers' individual behaviour and interactions with other drivers.

Type	*Indicator*	*Data type*	*Applicability in video analysis*
I	Driver's gender	value	poor
	Driver's age	value	poor
II	Vehicle lateral position	cont.	X-Y
	Lane deviation	yes/no	X-Y
	Lane departure corrected by steering	yes/no	X-Y
	Intersection approach speed	value	X-Y
	Lateral acceleration	cont.	X-Y
	Longitudinal acceleration	cont.	X-Y
	Steering	cont.	poor
	Time-to-Lane crossing (TLC)	cont.	X-Y
	Brake response time (BRT)	value	X-Y compl.
	Stop time (ST)	value	X-Y
	Stopping distance	value	X-Y
	Stopping accuracy (SA)	yes/no	X-Y
	Distracter response time (DRT)	value	poor
	Distracter response accuracy	yes/no	poor
	Gap acceptance	yes/no	X-Y
	"Red runner"	yes/no	X-Y compl.
	Arriving in dilemma zone	yes/no	X-Y compl.
	Compliance with STOP sign	yes/no	X-Y compl.
III	Collision course	yes/no	X-Y
	Time-to-Collision (TTC)	cont.	X-Y
	Minimal Time-to-Collision (TTC_{min})	value	X-Y
	Time-to-Collision with constant acceleration (TTCA)	cont.	X-Y
	Time-to-Collision with constant angular velocity	cont.	X-Y
	Time Exposed Time-to-Collision (TET)	value	X-Y
	Time Integrated Time-to-Collision (TIT)	value	X-Y
	Time-to-Collision for start of braking (TTC_{br})	value	X-Y
	Time-to-Collision for driver's first head movement (TTC_{1st})	value	X-Y
	Time-to-Collision when reaching the intersection stop line (TTI)	value	X-Y
	Enhanced Time-to-Collision (ETTC)	cont.	X-Y
	Inversed Time-to-Collision (1/TTC)	cont.	X-Y
	Post-Encroachment-Time (PET)	value	X-Y
	Gap Time	value	X-Y
	Encroachment Time	value	X-Y
	Initially Attempted PET (IAPT Time headway)	value	X-Y
	Time headway	value	X-Y
	Time gap	value	X-Y
	"Free vehicle"	yes/no	X-Y
	Compliance with yielding rules	yes/no	X-Y
IV	Number of "red runners"	value	X-Y compl.
	Level-of-Service	value	X-Y compl.

Várhelyi, 1998, summarised previous research on driver behaviour at non-signalised pedestrian crossings. Among the indicators found to be good explanatory variables were the pedestrian's position at the crossing (approaching the crosswalk, staying at the curb, ready to step out or crossing the road), *size of the pedestrian group*, *vehicle speed*, *size of the vehicle platoon* and the city size. Some other variables (*road width*, *presence of refuge*, *pedestrian age* and *gender*, whether pedestrian is *pushing a baby carriage* or *a bicycle*) were not found to be significant. The communication signals given by a pedestrian to the driver (i.e. *eye contact with the driver*, *putting foot on a carriageway*, *making hand signs*) were found to affect the driver's yielding behaviour. It was also noted that the "*free vehicles*" were of central importance for safety. The focus of the study was on the *driver's* choice of *speed* when approaching the zebra crossing. The interactive situations between a driver and pedestrian(s) were classified as follows: *pedestrian presence* – a pedestrian is approaching or crossing the zebra crossing when the approaching car is within 70 m from the crossing; encounter – a situation when a pedestrian and vehicle theoretically can arrive at the same point at the same time (i.e., they are on a collision course); conflict – a critical encounter before the possible collision with a time margin below the limit defined by the Swedish Traffic Conflict Technique (see a separate section on Traffic Conflicts below). To describe the driver's behaviour a *time-to-zebra* (TTZ) parameter was used, defined as "the distance to the zebra crossing divided by the speed at any given moment in time". The TTZ_{arr} is "the time left for the car to reach the zebra crossing at the moment the pedestrian arrives at the curb. The TTZ_{brake} value shows how much time there is left to the zebra crossing at the moment the braking starts, and gives an indication of the "driver's readiness to stop before the zebra crossing". The "pedestrian presence" situations are resolved by a pedestrian passing either *in front of the vehicle* or *behind* it. The first type of situation was further classified into "*no braking*", where the pedestrian passes without influencing the vehicle's speed, "*provoked braking*", when the driver is forced to brake to avoid collision or very low time gap between the pedestrian leaving and the driver arriving at the potential collision point and "*ideal interactions*" – when "the approaching car brakes on the driver's own initiative in order to give way to the pedestrian".

An interesting application of semi-automated video analysis is found in *Andersson, 2000*, where pedestrian behaviour is studied at two non-signalised urban intersections. The indicators used are *stop frequency* (number of stops per passage), *stop location*, the *stop length* (i.e. waiting time), the *crossing angle*, *location* and *speed on the crosswalk*. Table 4 summarizes the indicators mentioned above.

4.4. Cyclist behaviour and driver-cyclist interactions

Studying the safety and other effects of cycling lanes *Nilsson, 2003*, looked at the *distance* cyclists and drivers kept *to the kerb* or *parked cars* on the right, if the cyclist was moving in the *prescribed direction*, *on a cycle lane* or *a sidewalk*, and if a *group of cyclists* was moving *in a row* or *side by side*. Cyclists *with a child in a back seat* were marked separately. Other studies reviewed in this work also reported changes in *cyclists' speed*, *frequency of red-light violations* and the way the *left turn manoeuvre*

was performed (using a crosswalk, cycle crossing or cycling diagonally through the intersection).

Table 4. Indicators describing drivers' and pedestrians' individual behaviour and interactions.

Type	*Indicator*	*Data type*	*Applicability in video analysis*
I	Pedestrian age (child/adult)	value	poor
	Pedestrian gender	value	poor
	Pedestrian pushing a bicycle or baby carriage	yes/no	visual control
	Size of pedestrian group	value	X-Y
	Vehicle type	value	X-Y
	Vehicle platoon size	value	X-Y
II	Pedestrian approaches the crosswalk	yes/no	X-Y
	Pedestrian slows down at the curb	yes/no	X-Y
	Pedestrian stops at the curb	yes/no	X-Y
	Pedestrian presses the button to operate the traffic light	yes/no	X-Y compl./ visual control
	Pedestrian checks for approaching traffic (head movements)	yes/no	visual control/ poor
	Pedestrian steps off the sidewalk	yes/no	X-Y
	Pedestrian crosses within/outside the marked crosswalk	yes/no	X-Y
	Crossing location	cont.	X-Y
	Pedestrian crossing angle	value	X-Y
	Pedestrian walks on the crosswalk (rather than runs)	yes/no	X-Y
	Pedestrian speed on the crosswalk	cont.	X-Y
	Pedestrian is in the middle of the crossing	yes/no	X-Y
	Parent gives an oral instruction to a child	yes/no	visual control
	Parent holds the child's hand	yes/no	visual control
	Parent gives the child a chance to press the button	yes/no	visual control
	Pedestrian waits for green/violates red light	yes/no	X-Y compl.
	"Conflicting" red light violation (when conflicting flow has green)	yes/no	X-Y compl.
	Pedestrian waiting time	value	X-Y
	Pedestrian expected delay	value	X-Y compl.
	Pedestrian realized delay	value	X-Y
	Pedestrian stop frequency	value	X-Y
	Pedestrian stop location	value	X-Y
	Vehicle speed	cont.	X-Y
	Vehicle speed at a control point before the crossing	value	X-Y
	Vehicle speed when reaching the crosswalk	value	X-Y
	Vehicle journey time	value	X-Y
	Vehicle queue length	value	X-Y

III	Pedestrian is present	yes/no	X-Y
	Pedestrian communicates by putting foot on a carriageway	yes/no	visual control
	Pedestrian communicates through eye contact with the driver	yes/no	poor
	Pedestrian communicates by making hand signs	yes/no	visual control
	Pedestrian crosses before/after the vehicle	yes/no	X-Y
	Vehicle is present (oncoming traffic)	yes/no	X-Y
	Vehicle approaching on a lane close to/far from the pedestrian	yes/no	X-Y
	Encounter between pedestrian and vehicle	yes/no	X-Y
	"Free vehicle"	yes/no	X-Y
	Time-to-zebra (TTZ)	cont.	X-Y
	Time-to-zebra when the pedestrian arrives at the curb (TTZ_{arr})	value	X-Y
	Time-to-zebra when vehicle braking starts (TTZ_{brake})	value	X-Y
	Driver yields to pedestrian	yes/no	X-Y
	Vehicle brakes/does not brake	yes/no	X-Y
	Provoked vehicle braking	yes/no	X-Y
	Conflict in driver-pedestrian interaction	yes/no	X-Y
	"Ideal" driver-pedestrian interactions	yes/no	X-Y
IV	Proportion of pedestrians arriving on green	value	X-Y
	Proportion of pedestrians starting and completing the crossing on green	value	X-Y
	Proportion of pedestrians experiencing long delays	value	X-Y
	Proportion of pedestrians arriving on red and violating the red light	value	X-Y compl.
	Conflicts to flow ratio	value	X-Y

To describe the effects of a new traffic regulation concerning yielding at bicycle crossings, *Räsänen et al., 1999*, used the cyclist and vehicle speed and the *cyclist's head movements* as safety indicators. Cyclists approaching the crossing and *looking to the left, right* or *in both directions* and the approximate *age* of the cyclists (with 10-year interval) were recorded. The speed profiles were produced for *free* (i.e. crossing alone or being first in the queue) cars and cyclists only. The cases when a driver and a cyclist were on a collision course, or the situation developed into a conflict, were analysed separately. The *yielding behaviour* of the drivers was also registered.

Summala et al., 1996 studied drivers' visual search when approaching intersections with a cycle path. The main parameter registered was the *direction of the driver's head movement*. A concept of *critical area* was introduced, i.e. the area where the driver has to make a decision on whether to brake or continue driving, which depends on the vehicle speed and lies between the remotest point at which the driver starts to see the cyclist moving on a collision course, and the closest position where a start of braking still makes it possible to stop before the cycle path.

Räsänen & Summala, 2000 examined drivers' behaviour at roundabouts when a cyclist was present. Similar indicators to the ones above were used – the vehicle speed at the cycle crossing for different car situations – free car, *presence of other vehicles at the roundabout* and *presence of a cyclist approaching* from the left and from the right. Head movement, used as an indicator of a driver's visual search direction, was estimated as a deviation in degrees from the central line. Whether a driver yielded to a cyclist or not was also registered.

Table 5 summarises the above-mentioned indicators of cyclist behaviour and interactions with other road users.

Table 5. Indicators describing drivers' and cyclists' individual behaviour and interactions.

Type	*Indicator*	*Data type*	*Applicability in video analysis*
I	Cyclist's age	value	poor
	Cyclist has a child in the rear seat	yes/no	poor
II	Cyclist moves in prescribed direction	yes/no	X-Y
	Cyclist moves on cycle lane/sidewalk	yes/no	X-Y
	Group of cyclists moves in single file/side by side	yes/no	X-Y
	Cyclist makes a left turn (choice of trajectory)	value	X-Y
	Cyclist looks left/right/both (head movement)	value	visual control
	Cyclist violates the red light	yes/no	X-Y compl.
	Cyclist's speed	cont.	X-Y
	Cyclist's side distance from the kerb (parked vehicles) on the right	cont.	X-Y
	Vehicle enters the "critical area"	yes/no	X-Y
	Driver's head direction	value	poor
	Vehicle speed	cont.	X-Y
	Vehicle side distance from the kerb (parked vehicles) on the right	cont.	X-Y
III	Cyclist is present	yes/no	X-Y
	Vehicle is present	yes/no	X-Y
	"Free" cyclist	yes/no	X-Y
	"Free" vehicle	yes/no	X-Y
	Driver yields to a cyclist	yes/no	X-Y
	Collision course	yes/no	X-Y
	Conflict in cyclist-driver interactions	yes/no	X-Y

4.5. Traffic conflicts

Traffic Conflict Techniques are methods to estimate traffic safety based on observation of events that are not as severe as accidents, but similar to them in terms of the mechanisms by which they develop (*Hydén, 1987*). Most often the main concepts and definitions used in these techniques are equally applicable in vehicle-vehicle, vehicle-pedestrian or vehicle-cyclist studies.

Amundsen & Hydén, 1977, define a *conflict* as a situation where two or more road users are on collision course, and if their movements remain unchanged, they will collide. To classify the conflict severity, the Swedish Traffic Conflict Technique distinguishes between serious and non-serious conflicts using the parameters *Conflicting Speed* and *Time-to-Accident.* Time-to-Accident (TA) is a specific Time-to-Collision value defined as the time from the moment one of the road users starts an *evasive action* to the collision that would occur if they continued with unchanged speed and direction. The Conflicting Speed is the speed of the road user taking evasive action at the moment just before the start of the evasive action (*Svensson, 1998, Hydén, 1987*). TA is calculated by dividing the *distance to the collision point* by the Conflicting Speed.

Some conflict techniques use slightly different parameters to define a conflict, for example TTC_{min} or PET (e.g. *van der Horst & Kraay, 1986, Cooper, 1983*). However, quite a good agreement between the results of conflict studies performed by different techniques was found in an international calibration study (*Grayson, 1984*).

Further classification of the conflicts by their severity can be made by the *Deceleration Rate* (DR) at which a vehicle must brake to avoid a collision (*Gettman & Head, 2003*). *Malkhamah et al., 2005*) stated that vehicle deceleration rate itself was a valid indicator for conflicts and used it for conflict detection at a Pelican pedestrian crossing. *Nygård, 1999*, found that a derivative of the deceleration, *jerk*, is significantly different in serious conflicts as compared to other types of braking.

Table 6 summarizes the above-mentioned indicators.

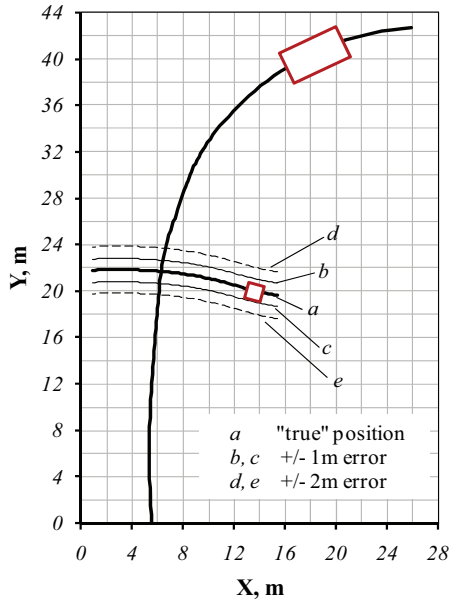
Table 6. Indicators describing conflict development and severity.

Type	*Indicator*	*Data type*	*Applicability in video analysis*
II	Distance to collision point	value	X-Y
	Conflicting Speed	value	X-Y
	Deceleration Rate (DR)	value	X-Y
	Jerk	value	X-Y
III	Collision course	yes/no	X-Y
	Evasive action	yes/no	X-Y compl.
	Time-to-Accident (TA)	value	X-Y compl
	Minimal Time-to-Collision (TTC_{min})	value	X-Y
	Post-Encroachment Time (PET)	value	X-Y
	Conflict	yes/no	X-Y

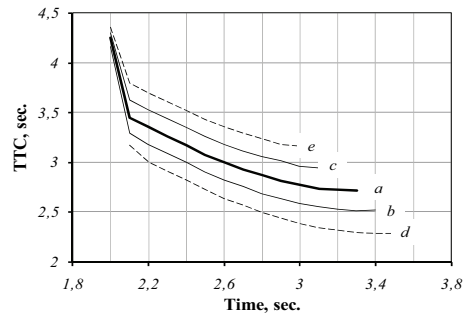
5. ACCURACY OF THE COMPLEX INDICATORS

Even though the results in Tables 3-6 suggest that many indicators may be expressed through "raw" data like road users' position and speed, the next challenge in implementing them in a video analysis system is that the indicator accuracy depends highly on how accurately the speed and position are estimated. This is illustrated in an example in Figure 4, which shows calculations of Time-to-Collision and Post-Encroachment Time for an interaction between a pedestrian and a car. To begin with, the two road users appear on a collision course, but then the car driver slows down and lets the pedestrian pass first. Calculations are performed for the "true" pedestrian trajectory (*a* in Figure 4*a*) and several trajectories with introduced errors in positions generated by shifting the "true" trajectory by 1 and 2 meters (*b-e*, Figure 4*a*).

Figure 4*b* shows the Time-to-Collision curves calculated for the different pedestrian trajectories. The table in Figure 4*c* presents the minimal Time-to-Collision values (TTC_{min}) and Post-Encroachment Time (PET) for the different pedestrian trajectories.



a)



b)

Trajectory	TTC_{min}, sec.	PET, sec.
a ("true")	2.7	1.3
b	2.5	1.2
c	2.9	1.5
d	2.3	0.9
e	3.2	1.8

c)

Figure 4. The effects of position accuracy on Time-to-Collision (TTC) and Post-Encroachment Time (PET) estimates: a) vehicle and pedestrian trajectories; b) TTC profiles calculated for different pedestrian trajectories; c) calculated TTC_{min} and PET.

The error in position has an effect on both the TTC and PET values, as well as on how long the road users are considered to be on a collision course. The span of TTC_{min} is between 2.3 and 3.2 seconds, while PET varies between 0.9 and 1.8 seconds. Obviously, if these values are used to calculate indicators like Time Exposed or Time Integrated TTC, the difference in results will be even more striking.

6. DISCUSSION AND CONCLUSIONS

Automated video analysis is a set of techniques that can detect moving objects in video recordings and produce their trajectories and to some extent other parameters like speed, size, type, etc. At the current stage, the technology has many limitations (problems with camera installations, limited study area, long time required to process the data, low detection rate and high number of false detections, limited accuracy of the position and speed estimates, need for extensive manual control of the results, etc.), but there is also a clear trend in quality improvements and increase in the degree of automation of the tasks as the techniques become more complex and sophisticated. The computational power of available computer hardware is constantly increasing, too.

Out of 119 unique indicators in 45 articles (Tables 3-6), 98 (86%) indicators currently used in traffic behaviour research can be calculated using the results of video analysis only or with some additional input. These indicators are expressed though parameters like position, speed, direction, size, type of road user, i.e., the direct

output from the video analysis algorithms. The indicators that will be hard to extract from video data describe less visible parameters such as personal characteristics of road users (like age and sex), and actions like head, eye and hand movements and eye contact.

The accuracy of the position and speed estimation done from video has a great effect on the accuracy of the indicators based on this data (Figure 4). This emphasizes the question of reliability of the video analysis as a measurement tool. A wide range of factors affect the accuracy of the video analysis results, e.g., camera location and view angle, resolution of video images and frame rate, atmospheric and light conditions, traffic conditions and flow composition, etc. Obviously, at each site, or even at the same site but during different time periods, this combination of these factors will be unique and thus the detection quality will never be the same. If the errors in detection occur just by chance, they can be compensated for simply by extending the observation period until the number of correct detections is sufficient. However, if the detection quality depends on some condition (e.g. traffic intensity) which, in turn, is related to the studied quality (e.g. safety), the results become systematically biased. It is important therefore to perform extensive quality tests of video analysis techniques parallel with some other detection method with established reliability (e.g. manual observations) in various conditions to ensure the reliability of the video analysis system.

Producing highly accurate trajectories for all road users requires very intensive calculations, which might not be feasible during longer study periods. A possible solution may be to perform detection in two steps, i.e., first filter potentially interesting situations using not very accurate trajectories and low threshold values, and then produce more accurate road user trajectories for these situations only. The risk, however, is that the low threshold may result in too many detections that will still be hard to process.

The indicator review also shows that indicators of the "yes/no"- and single “value”-type dominate when describing road users’ individual behaviour. Some of the indicators can only be described in this way; for example, the question of whether a pedestrian crosses a street before or after a car. For many of the other indicators the data type can easily be modified to become continuous when an instrument capable of measuring the parameter continuously is available (e.g. vehicle speed can be measured at a certain fixed point, but also as a speed profile over time, *Laureshyn et al., 2009*). The use of continuous rather than point indicators implies that more data is collected, which may be analyzed in a more advanced way and new, more advanced or integrated parameters may be introduced. For example, the acceleration profile and the second derivative of speed – jerk – can be calculated from the speed profile, but not from a point speed value.

It is not the intention of this paper to judge which of the indicators are valid representatives for safety or comfort. Surprisingly, the validity issue hardly gets any attention in any of the reviewed publications especially since, except for a very few thoroughly studied indicators (e.g. serious conflicts and speed), the relation between an indicator and the described quality is far too often based on assumptions and

common sense. Establishing the validity is quite a challenge as it requires large data samples representing different locations and conditions, and with conventional techniques this extensive data collection is hardly feasible. From this perspective, video analysis offers a unique opportunity to collect larger datasets and thus contribute to the validation of existing indicators as well as newly proposed ones.

The prospect of being able to collect much more data, and of different types than is possible today, is double-edged, though. It will be possible to conduct more sophisticated data analysis and come up with more complex and integrated indicators. On the other hand, the validity issue of the selected indicators will become even more important. The data to be collected should not only be collected on the basis of "it is so easy to collect"; it must also be selected on the basis of being a good representative of the qualities of interest. A great advantage could be that extended observation periods would also allow the registration of direct safety measures like accidents.

Further automation of the analysis requires strict definitions of all calculation algorithms, indicators and their critical values. Even if one considers this a limitation or lack of flexibility, it may also be regarded as a means of improving the objectiveness of the measurements and making the results of different studies and sites more comparable.

7. REFERENCES

Allen, B. L., B. T. Shin, P. J. Cooper (1977). Analysis of traffic conflicts and collisions. Department of Civil Engineering, McMaster University.

Amundsen, F. N., C. Hydén (1977). Proceedings of the first international traffic conflicts technique workshop. Institute of Transport Economics, Oslo.

Andersson, J. (2000). Image processing for analysis of road user behaviour – a trajectory based solution. Lund Institute of Technology, Department of Technology and Society, Traffic Engineering. Bulletin 212.

Archer, J. (2005). Indicators for traffic safety assessment and prediction and their application in micro-simulation modelling: a study of urban and suburban intersections. Doctoral thesis, Royal Institute of Technology, Stockholm, Department of Infrastructure.

Ardö, H. (2009). Multi-target tracking using on-line viterbi optimisation and stochastic modelling. Doctoral thesis, Lund University, Faculty of Engineering LTH, Centre for Mathematical Science.

Ardö, H., K. Åström (2008). Bayesian formulation of image patch matching using cross-correlation. Computer Vision and Pattern Recognition. Proceedings of IEEE Computer Society Conference on CVPR 08'.

Atev, S., H. Arumugam, O. Masoud, R. Janardan, N. P. Papanikolopoulos (2005). A vision-based approach to collision prediction at traffic intersections. IEEE transactions on intelligent transportation systems 6 (4), pp. 416-423.

Autoscope® (2009). Video detection solutions to improve traffic flow and roadway safety. www.autoscope.com *Acc.* 9 February 2009.

Berclaz, J., F. Fleuret, P. Fua (2006). Robust people tracking with global trajectory optimization. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, USA.

Björklund, G. M., L. Åberg (2005). Driver behaviour in intersections: formal and informal traffic rules. Transportation Research Part F 8, pp. 239–253.

Cognimatics (2009). TRUEVIEW People CounterTM. http://www.cognimatics.com/prodTVPC.php *Acc.* 13 Mar 2009.

Coifman, B., D. Beymer, P. McLauchlan, J. Malik (1998). A real time computer vision system for vehicle tracking and traffic surveillance. Transportation Research Part C 6, pp. 271-288.

Cooper, P. J. (1983). Experience with traffic conflicts in Canada with emphasis on post encroachment time techniques. Springer-Verlag, Heidelberg.
Dingus, T. A., V. L. Neale, S. G. Klauer, A. D. Petersen, R. J. Carroll (2006). The development of a naturalistic data collection to perform critical incident analysis: an investigation of safety and fatigue issues in long-haul trucking. Accident Analysis and Prevention 38, pp. 1127-1136.
Elvik, R. (1997). Evaluations of road accident blackspot treatment: a case of the iron law of evaluation studies? Accident Analysis & Prevention 29 (2), pp. 191-199.
Englund, A., N. P. Gregersen, C. Hydén, P. Lövsund, L. Åberg (1998). Trafiksäkerhet – en kunskapsöversikt (*in Swedish*) Traffic safety - a knowledge review. KFB, Studentlitteratur, Lund.
Evans, L. (1991). Traffic Safety and the Driver. Van Nostrand Reinhold, New York.
Fernandez-Caballero, A., F. J. Gomez, J. Lopez-Lopez (2007). Road-traffic monotoring by knowledge-driven static and dynamic image analysis. Expert Systems with Applications *35* (3), pp. *701-719*.
Gettman, D., L. Head (2003). Surrogate safety measures from traffic simulation models. U. S. Department of Transportation, Federal Highway Administration. Report No. FHWA-RD-03-050.
Grayson, G. B. (1984). The Malmö study: a calibration of traffic conflict techniques. SWOV. Report R-84-12.
Hagring, O. (2000). Framkomlighet i korsningar utan trafiksignaler, en litteraturöversikt (*in Swedish*) Traffic efficiency of non-signalised intersections, a literature overview. Lunds Tekniska Högskola, Institutionen för Teknik och Samhälle, Lunds Universitet. Bulletin 190.
Hakkert, A., V. Gitelman, E. Ben-Shabat (2002). An evaluation of crosswalk warning systems: effects on pedestrian and vehicle behaviour. Transportation Research Part F 5, pp. 275-292.
Hancock, P. A., M. Lesch, L. Simmons (2003). The distraction effects of phone use during a crucial driving manoeuvre. Accident Analysis and Prevention 35, pp. 501-514.
Harré, N., W. Wrapson (2004). The evaluation of a central-city pedestrian safety campaign. Transportation Research Part F 7, pp. 167-179.
Hayward, J. C. (1971). Near misses as a measure of safety at urban intersections. Doctoral thesis, The Pensilvania State University, Department of Civil Engineering.
Hitachi (2008). ITS21, Traffic Surveillance System. http://www.hitachi.co.jp/en/products/its/product/service/2004030_12381.html *Acc.* 9 Feb 2009.
Hydén, C. (1987). The development of a method for traffic safety evaluation: the Swedish traffic conflict technique. Doctoral thesis, Lund University, Department of Traffic Planning and Engineering.
Isard, M., J. MacCormick (2001). BraMBLe: A bayesian multiple-blob tracker. IEEE International Conference on Computer Vision.
Ji, X., Z. Wei, Y. Feng (2006). Effective vehicle detection technique for traffic surveillance systems. Journal of Visual Communication and Image Representation 17, pp. 647-658.
Kiefer, R. J., D. J. LeBlanc, C. A. Flannagan (2005). Developing an inverse time-to-collision crash alert timing approach based on drivers' last-second braking and steering judgments. Accident Analysis and Prevention 37, pp. 295-303.
Koller, D., K. Daniilidis, T. Thórhallson, H.-H. Nagel (1992). Model-based object tracking in traffic scenes *in* G. Sandini (Ed.) Computer Vision - ECCV'92. Springer, Berlin/Heidelberg.
Köll, H., M. Bader, K. W. Axhausen (2004). Driver behaviour during flashing green before amber: a comparative study. Accident Analysis and Prevention 37, pp. 273-280.
Laureshyn, A., H. Ardö (2006). Automated video analysis as a tool for analysing road user behaviour. ITS World Congress, London, UK.
Laureshyn, A., H. Ardö, T. Jonsson, Å. Svensson (2009). Application of automated video analysis for behavioural studies: concept and experience. IET Intelligent Transport Systems 3 (3), pp. 345-357.
Laureshyn, A., K. Åström, K. Brundell-Freij (2009). From speed profile data to analysis of behaviour: classification by pattern recognition techniques. IATSS Research 33 (2), pp. 88-98.
Leibe, B., K. Schindler, N. Cornelis, L. V. Gool (2008). Coupled object detection and tracking from static cameras and moving vehicles. Pattern Analysis and Machine Intelligence, IEEE Transactions 30 (10), pp. 1683-1698.

Malkhamah, S., M. Tight, F. Montgomery (2005). The development of an automatic method of safety monitoring at Pelican crossings. Accident Analysis and Prevention 37, pp. 938-946.
Martinez, K. L. H., B. E. Porter (2006). Characterizing red light runners following implementation of a photo enforcement program. Accident Analysis and Prevention 38, pp. 862–870.
Messelodi, S., C. M. Modena (2005). A computer vision system for traffic accident risk measurement: a case study. Instituto Trentino di Cultura (ITC). Technical Report T05-06-07.
Messelodi, S., C. M. Modena, M. Zanin (2004). A computer vision system for the detection of vehicles at urbam road intersections. Instituto Trentino di Cultura (ITC). Technical Report T04-02-07.
Minderhoud, M. M., P. H. L. Bovy (2001). Extended time-to-collision measures for road traffic safety assessment. Accident Analysis and Prevention 33, pp. 89-97.
Moon, J. Y. J., F. Coleman III (2003). Dynamic dilemma zone based on driver behaviour and car following model at highway-rail intersections. Transportation Research Part B 37, pp. 323-344.
Nilsson, A. (2003). Utvärdering av cykelfälts effekter på cyklisters säkerhet och cykelns konkurrenskraft mot bil (*in Swedish*) Validation of the effects of bicycle lanes on bicyclists' safety and bicycle competition power against a car. Doctoral thesis, Lunds Universitet, Lund, Institutionen för Teknik och Samhälle.
Nygård, M. (1999). A method for analysing traffic safety with help of speed profiles. Master thesis, Tampere University of Technology, Department of Civil Engineering.
Parkhurst, D. (2006). Using digital video analysis to monitor driver behaviour at intersections. Center for Transportation Research and Education, Iowa State University. CTRE Project 05-214.
Räsänen, M. (2005). Effects of a rumble strip barrier line on lane keeping in a curve. Accident Analysis & Prevention 37 (3), pp. 575-581.
Räsänen, M., I. Koivisto, H. Summala (1999). Car driver and bicyclist behaviour at bicycle crossings under different priority regulations. Journal of Safety Research 30, pp. 67-77.
Räsänen, M., H. Summala (2000). Car drivers' adjustment to cyclist at roundabouts. Transportation Human Factors 2, pp. 1-17.
Sakamoto, Y., K. Kajitani, T. Naito, S. Kamijo (2006). Development of the image processing vehicle detector for intersections. ITS World Congress, London.
Sakshaug, L., Å. Svensson, C. Hydén, A. Laureshyn (2009). Bicyclists in roundabouts – different design solutions. Accident Analysis & Prevention (***submitted***).
Saunier, N., T. Sayed (2007). Automated analysis of road safety with video data. Transportation Research Record: Journal of the Transportation Research Board 2019, pp. 57-64.
Schoenemann, T., D. Cremers (2008). Globally optimal shape-based tracking in real-time. IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, Alaska.
Shi, J., C. Tomasi (1994). Good features to track. Computer Vision and Pattern Recognition. Proceedings of IEEE Computer Society Conference on CVPR, pp. 593-600.
Song, X., R. Nevatia (2007). Detection and tracking of moving vehicles in crowded scenes. IEEE Workshop on Motion and Video Computing.
Springboard (2009). Customer counting and performance measurement. http://www.spring-board.info *Acc.* 13 Mar 2009.
Summala, H. (1980). How does it change safety margins if overtaking is prohibited: a pilot study. Accident Analysis and Prevention 12, pp. 95-103.
Summala, H., E. Pasanen, M. Räsänen, J. Sievänen (1996). Bicycle accidents and drivers' visual search at left and right turns. Accident Analysis and Prevention 28, pp. 147-153.
Suzuki, K., H. Nakamura (2006). TrafficAnalyzer - the integrated video image processing system for traffic flow analysi. ITS World Congress, London, UK.
Svensson, Å. (1998). A method for analysing the traffic process in a safety perspective. Doctoral thesis, University of Lund, Lund Institute of Technology, Department of Traffic Planning and Engineering.
Tan, T. N., G. D. Sallivan, K. D. Baker (1998). Model-based localisation and recognition of road vehicles. International Journal of Computer Vision 27 (1), pp. 525.
Traficon® (2008). Video for traffic data acquisition, automatic incident and presence detection. www.traficon.com *Acc.* 9 Feb 2009.
TRB (2000). Highway Capacity Manual 3rd edition. US Transportation Research Board. Special Report 209.

van der Horst, R. (1990). A time-based analysis of road user behaviour in normal and critical encounters. Doctoral thesis, Delft University of Technology.
van der Horst, R., R. J. Kraay (1986). The Dutch conflict technique – DOCTOR. ICTCT Workshop, Budapest.
van Winsum, W., K. A. Brookhuis, D. de Waard (2000). A comparison of different ways to approximate time-to-lane crossing (TLC) during car driving. Accident Analysis and Prevention 32, pp. 47-56.
van Winsum, W., D. de Waard, K. A. Brookhuis (1999). Lane change manoeuvres and safety margins. Transportation Research Part F 2, pp. 139-149.
Várhelyi, A. (1998). Driver's speed behaviour at a zebra crossing: a case study. Accident Analysis and Prevention 30 (6), pp. 731-743.
VisioWay (2007). Video detection systems. www.visioway.com *Acc.* 9 Feb 2009.
Vogel, K. (2002). What characterizes a "free vehicle" in an urban area? Transportation Research Part F 5, pp. 15-29.
Vogel, K. (2003). A comparison of headway and time to collision as safety indicators. Accident Analysis and Prevention 38, pp. 427-433.
Zeedyk, M. S., L. Kelly (2003). Behavioural observations of adult-child pairs at pedestrian crossings. Accident Analysis and Prevention 35, pp. 771-776.
Zhao, T., R. Nevatia (2004). Tracking multiple humans in complex situations. Pattern Analysis and Machine Intelligence, IEEE Transactions 26 (9), pp. 1208-1221.
Åström, K., J. Karlsson, O. Enquist, A. Ericsson, F. Kahl (2007). Automatic Feature Point Correspondences and Shape Analysis with Missing Data and Outliers *in* B. K. Ersböll and K. S. Stenstrup (Eds.). Lecture Notes in Computer Science. Springer.

Paper II

Submitted to "Accident Analysis and Prevention" on 3rd December 2009

EVALUATION OF TRAFFIC SAFETY, BASED ON MICRO-LEVEL BEHAVIOURAL DATA: THEORETICAL FRAMEWORK AND FIRST IMPLEMENTATION

Aliaksei Laureshyn, Åse Svensson, Christer Hydén

Traffic and Roads, Department of Technology and Society,
Faculty of Engineering LTH, Lund University

Box 118, 22100 Lund, Sweden Tel. +46 46 222 91 25 Fax +46 46 12 32 72

Abstract. A traffic encounter between individual road users is a process of continuous interplay over time and space and may be seen as an elementary event with the potential to develop into an accident. This paper proposes a framework for organising all traffic encounters into a severity hierarchy based on some operational severity measure. A severity hierarchy provides a description of the safety situation and trade-off between safety and efficiency in the traffic system. As a first approach to studying the encounter process, a set of indicators is proposed to describe an encounter. These indicators allow for a continuous description even if the relationship between the road users changes during the process (e.g. when they are on a collision course or leave it). Automated video analysis is suggested as a tool that will allow data collection for validation of the proposed theories.

Keywords: automated video analysis, encounter process, road user behaviour, severity hierarchy, traffic safety indicators

1. INTRODUCTION

Traditionally, road traffic safety analysis has relied mostly on accident statistics as the main data source. Over the years, however, numerous problems associated with accident data have been discussed. To sum up, the following aspects are of importance: i) compared to other events in traffic, accidents are very exceptional in the sense that they are the results of a series of unhappy realisations of many small probabilities; ii) accidents are rare events, making it troublesome to base traffic safety analyses at individual sites on accidents only; iii) not all accidents are reported and the level of underreporting depends on the accident's severity and types of road users involved; iv) information on the behavioural aspects preceding the accident is seldom available.

There is a need to use some kind of surrogate and complement to accidents, i.e., traffic safety indicators, to increase the possibility of: i) evaluating traffic safety changes more efficiently and in a shorter time; ii) elaborating the relation between design elements and risk; iii) more thoroughly understanding the relationships between behaviour and risk; iv) a better understanding of the processes characterising the normal traffic and critical situations including accidents.

This paper is concerned with micro-level-behaviour indicators of traffic safety. Until recently, the use of such indicators has been quite limited. The existing traffic conflict techniques (*Hydén, 1987, van der Horst & Kraay, 1986, Asmussen, 1984*), even though very appealing from a theoretical point of view, have so far relied greatly on using human observers, a factor that limits the efficiency of data collection and the level of details it is possible to achieve. The relationships between many other behavioural indicators and safety have not been thoroughly validated and are far too often based on assumptions and common sense (*Laureshyn & Svensson, 2009*).

Automated video analysis is a rapidly developing technology that might provide a solution for effective behaviour data collection. Today's video analysis systems (e.g. *Laureshyn et al., 2009, Messelodi et al., 2004*) are already capable of detecting and tracking road users of various types, and there is a clear trend of increasing the studied area size, improving processing time and accuracy of the results. An optimistic, but quite reasonable, expectation is that in the relatively near future there will be a tool available to provide a detailed description of movements (i.e., co-ordinates related to time) of all road users within the studied area, for example an entire intersection. Such data has great potential for traffic safety analysis, but the practical methodology for it still needs to be developed. In the first place, this concerns the choice of safety indicators to be extracted from the data, the way they are to be analysed and how the results are to be interpreted.

The aim here is to propose a theoretical framework for the development of a method for traffic safety evaluation that utilises the detailed micro-level behavioural data provided by a video analysis system or similar tools. We also make a first attempt to develop a set of safety indicators that describe a continuous process of interaction between individual road users and relate the individual interactions to the general safety situation.

2. THEORETICAL FRAMEWORK

An encounter (a simultaneous arrival in a certain limited area) between two road users can be seen as an elementary event in the traffic process that has a potential to end up in a collision. *Hydén, 1987*, suggests the existence of some severity dimension common for all the events in traffic and proposes a model describing the relation between the events' severity and their frequency (Figure 1*a*). According to this model, the higher the severity (presented as the vertical position in the pyramid), the lower the frequency (the volume of the pyramid slice at this height) of the events.

The concept of severity requires some clarification. The severity of an accident is determined by the accident's consequences (e.g. number of deaths and injuries or total loss in monetary units). This definition is a bit problematic for encounters that do not end in a collision, as, strictly speaking, a near-miss with just a few centimetres between the vehicles and a completely controlled passage with sufficient safety margins have the same consequences (except for differences in the adrenaline level in the drivers' blood). Two aspects are to be considered here – the potential of an encounter to become an accident and severity of the consequences *if* this happens.

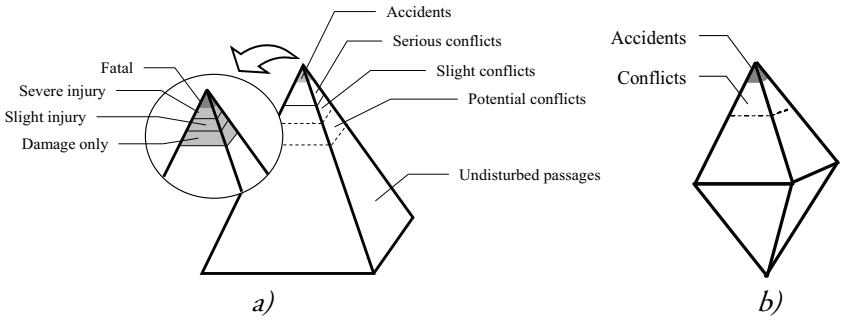


Figure 1. The relation between severity and frequency of elementary events in traffic: a) safety pyramid** (Hydén, 1987)**; b) diamond-shaped safety hierarchy** (*adopted from *Svensson, 1998)**.

The accident potential may be explained in the following way. Accidents are stochastic events. Even though one particular accident may be explained by a number of factors that led to it, it may also be considered as an unlucky coincidence that all these factors happened to be there at the same time. If some of the contributing factors had not been present, the accident might have been avoided. To put it another way, each encounter can develop into an accident if some new factors come up. In any case, a near-miss has less of a safety margin to endure an additional unlucky factor compared to a well-controlled passage; thus the severity of a near-miss is higher.

The severity of the accident's outcome is influenced by several factors:

- ***Type of road users.*** If all other variables are equal then: (i) unprotected road users (pedestrians, cyclists, moped drivers, motorcyclists) are likely to suffer more severe injuries than protected road users (people travelling in a car, bus, lorry); (ii) a person travelling in a vehicle of small mass is likely to suffer more severe injuries than a person in a vehicle of large mass; (iii) an elderly person is likely to suffer more severe injuries than a younger person (*Englund et al., 1998*).
- ***Collision angle.*** The road users' angle of approach before the collision may have many different patterns from head-on to rear-end. For collisions involving protected road users this implies different probabilities regarding the collision impact. Head-on collisions are less likely to produce severe injuries than perpendicular collisions (because of less protection provided by the vehicles sides), while rear-end collisions are less likely to produce severe injuries than the other collision types (*SIKA, 2008*). Presumably, the angle of approach does not have the same effect regarding collision impact when vulnerable road users are involved.
- ***Collision speed.*** Collisions at higher speeds produce more severe injuries than collisions at lower speeds due to a larger amount of kinetic energy released (*Carlsson, 2004*). There are indications that the relative speed of the involved road users is a more important variable compared to the absolute speed values.

We leave, for the moment, the question of whether the probability of a collision and the severity of the consequences are to be kept apart or not, and assume that the

severity dimension integrates both of them. When severity is assigned to the encounters, they can be placed in some kind of distribution similar to Hydén's pyramid (such distributions are called *severity hierarchies* in *Svensson, 1998*). The way the severity is defined determines the actual shape of the hierarchy. It is reasonable to assume that there is a *"true"* hierarchy which reflects the objective severity. Introducing various operational measures to describe the severity of an encounter may create many quite different hierarchies in which the same event will probably not be placed exactly on the same level.

If a severity hierarchy represents events at a particular site (e.g. an intersection), a more correct illustration will not be a pyramid but a diamond (Figure 1*b*). The least severe events in traffic are quite rare when a road user is completely undisturbed by other road users. The majority of the encounters are of "medium severity", i.e. road users have to adjust their actions to the other road users, but in a well-controlled manner that characterises the "normal" traffic process. *Svensson, 1998*, also argues for the doubled peak shape of the distribution, but she limits the events included in the hierarchy to only those with a collision course (i.e., at some point the road users will collide if they continue with unchanged speed and path).

Severity hierarchy gives a much better understanding of the situation from a safety point of view compared to accidents that only represent the very top of the distribution. The important question is how the frequency of events in different severity levels is to be interpreted. A robust relation between the frequency of serious conflicts and the actual number of police-reported accidents has been found (*Hydén, 1987*). Findings in *Svensson, 1998*, suggest that the non-serious conflicts bear different information depending on how close to the serious conflicts they are located in the severity hierarchy. Events located just beneath the serious conflicts, i.e., events with fairly high severities, are characterised by closeness in time and space, thus still having a strong relation to safety. Studies at a non-signalised intersection showed an accumulation of interactions at these fairly high severities while there were no accidents or serious conflicts. Comparative studies at a signalised intersection revealed that interactions with fairly high severities did not seem to exist while accidents did. Svensson's interpretation is that these interactions at fairly high severities may be positive from a safety point of view, because they are frequent and severe enough to increase awareness, but not severe enough to result in accidents.

The hypothesis is thus that different hierarchies will have different degrees of accumulation of events allocated to different parts of the hierarchy. The shape and position of the part of the hierarchy where most events are located may possibly reflect the predominant road user behaviour. This behaviour may be interpreted as representing the road users' optimisation of their desires to keep a high mobility standard, and to maintain safety margins and comfort. This is very much in line with the theories explaining road users' behaviour, for example, risk homeostasis (*Wilde, 1994*) and zero-risk (*Näätänen & Summala, 1976*) theories, suggesting that road users' chosen safety margins also depend on considerations other than safety, such as the wish to maintain a certain speed or to conserve energy and comfort. *Svensson & Hydén, 2006*, and *Svensson, 1998*, also argue that the frequent interactions of fairly

high severities at locations without accidents or serious conflicts could reflect a situation where the mobility and safety desires of the involved road users are balanced.

3. OPERATIONAL DEFINITION OF SEVERITY

The way the severity of an encounter is to be defined in the best operational way is still an open issue. The following aspects are to be considered:

3.1. What measures are relevant?

The literature proposes indicators to reflect the two aspects of severity, the risk of collision and the collision consequences. As a measure of a collision risk, indicators describing proximity in space, proximity in time and intensity of a necessary evasive action can be mentioned (*Gettman & Head, 2003, Hydén, 1996, van der Horst, 1990, Hydén, 1987, Asmussen, 1984, Allen et al., 1977, Hayward, 1971*). The general problem is that if only one variable is used, just one side of the truth is reflected (for example, short distance between road users does not say much without information about their speeds). In this respect, the time-proximity indicators are a bit special since they integrate both proximity and speed. It is probably for this reason that many of the traffic conflict techniques base the conflict severity gradation on some kind of time-based severity measure (*Asmussen, 1984*). Still, it seems reasonable to complement time proximity measures with a speed indicator, as, for example, is done in the Swedish traffic conflict technique (*Hydén, 1987*).

The speed of the road users is also a relevant measure for estimation of the consequences. Other indicators that can reflect the severity consequences are road user type (or some alternative measure of "vulnerability), approaching angle, collision type, etc.

3.2. Relation to a collision course

The collision course at the end of an encounter is a pre-condition for a collision; without it a collision is not possible. Notwithstanding, even encounters without a collision course might have an accident potential, but some changes in spatial or temporal relation between the road users have to occur in order to reach collision course. Generally, these relations at a particular moment may be classified into three types (Figure 2):

Type A (collision course). Road users are on a collision course – they will collide if no evasive action is taken.

Type B (crossing course). Road users' planned paths overlap, but collision will be avoided as they pass the common spatial zone at different times. For collision to become possible a correction in time is needed; i.e., one or both road users have to change speed. In other words, the situation has to turn into an A-type situation first.

Type C (diverging course). Road users' paths do not overlap in any way. This occurs when road users have parallel or diverging courses. This does not mean, however, that the collision risk is completely zero as some (often very little) adjustment of the path by one or both road users may make their courses overlap and in certain conditions

also create a collision course. For example, a pedestrian walking close to the street curb is just a few steps away from cars driving at high speeds, and there is a risk that the situation will develop into a very severe one if he/she suddenly changes the direction of walking. It may be argued, though, that before a collision happens the situation must turn into an A-type situation (collision course), possibly via a B-type situation (crossing course) and then it can be described by methods developed for these situation types.

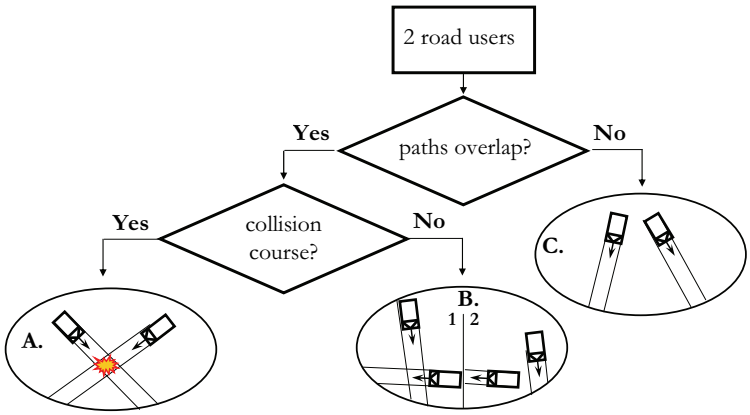


Figure 2. Classification of an encounter state based on spatial and temporal relations between the road users.

Even though these three types are theoretically different, they create a continuum in which the transfer between the types occurs smoothly. Since even minor changes in road users' trajectories (paths) and speed during an encounter may affect whether they appear on a collision course or not, or even if they pass a common spatial zone or not, the behaviour of the road users does not change abruptly at the moment of transfer. For example, *Svensson, 1998*, found that in situations when two vehicle drivers were about to miss each other by a very short time margin, their evasive behaviour was the same as if they were on a collision course. The most obvious explanation is that with short margins even a minor change in speed might put road users on a collision course, and therefore the situation is experienced as being high risk as well. The Dutch traffic conflict technique DOCTOR (*Kraay & van der Horst, 1985*) includes in its definition of a conflict both situations with a collision course and without a collision course, given that the time margin is small enough. This indicates that the measures used to describe an encounter also have to be flexible enough to allow smooth transfer in the description.

3.3. Encounter as a process

An encounter between two road users is a continuous process and the severity indicators to describe this process should also allow for a continuous description and not only for a certain moment during the process. As the encounter may go through different phases including moving on or not moving on a collision course, the severity indicators must allow for smooth transfers in the description. On the other hand, for

an encounter to be placed in a severity hierarchy, it has to be represented by only one value, which may be derived from a set of indicators all describing the severity of the process with regard to a certain aspect. The challenge is therefore to unify these indicators into a common severity measure that will make it possible to locate them in one common severity hierarchy.

4. A SET OF INDICATORS - A FIRST APPROXIMATION

We propose a set of indicators that continuously describe the process of an encounter and may be used to classify an encounter's severity. The purpose is to find the common severity measure that makes it possible to elaborate on a common severity hierarchy. The set includes several time-based indicators and the speed of the road users. Some of the problems discussed in the previous section are addressed, but a solution is not found for all of them. Therefore, this should be seen as a first approximation and further elaboration and validation of the approach are necessary.

4.1. Time-to-Collision

Time-to-Collision (TTC) is defined as the time required for two vehicles to collide if they continue at their present speed and along the same path (*Hayward, 1971*). Most often TTC is calculated by using the simple assumption that the road users' trajectories cross at a right angle or are parallel (Figure 3). For example, *van der Horst, 1990*, calculates TTC for the case of a right-angle approach by using the following equations:

$$\begin{cases} TTC = d_2/v_2, & if \quad d_1/v_1 < d_2/v_2 < (d_1 + l_1 + w_2)/v_1 \\ TTC = d_1/v_1, & if \quad d_2/v_2 < d_1/v_1 < (d_2 + l_2 + w_1)/v_2 \end{cases}, \qquad \textit{Eq. 1}$$

where d_1, d_2 are distances from the fronts of vehicles 1 and 2, respectively, to the area of intersection (Figure 3*a*);

l_1, l_2, w_1, w_2 are the lengths and widths of vehicles 1 and 2, respectively;

v_1, v_2 are the vehicle speeds.

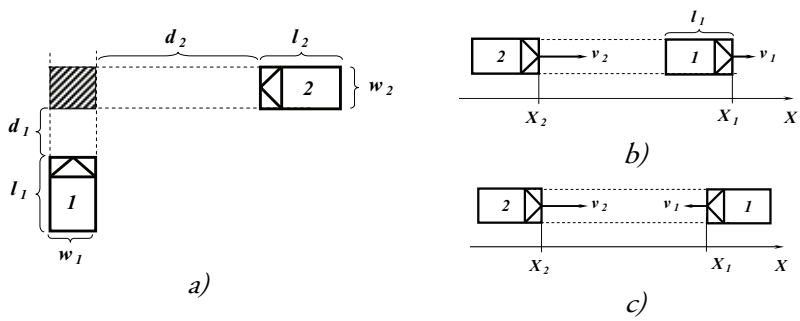


Figure 3. Calculation of TTC for perpendicular and parallel trajectories.

For the case of rear-end collision (Figure 3*b*) *Minderhoud & Bovy, 2001*, calculate TTC as:

$$TTC = \frac{X_1 - X_2 - l_1}{v_1 - v_2}, \quad \text{if} \quad v_2 > v_1, \qquad \textit{Eq. 2}$$

where X_1 and X_2 are the positions of vehicles 1 and 2, respectively.

For the case of a head-on collision (Figure 3*c*), the previous equation can be easily modified to:

$$TTC = \frac{X_1 - X_2}{v_1 + v_2}. \qquad \textit{Eq. 3}$$

In the general case, two vehicles can approach each other at any angle and, moreover, different collision types are possible for the same angle (Figure 4). After analysing all the possible collision types, it can be concluded that it is always a corner of one of the vehicles that meets a side of the other one. Since, in the general case, it is not known which corner meets which side, all possible combinations have to be analysed (i.e., 32 combinations assuming that road users have rectangular forms). The procedure for calculating TTC for a moving line section and a point (i.e., a side and a corner of road users) is given in the Appendix. The lowest TTC-value found among all the corner-side combinations is to be used, since it is the side and corner that will come into contact first in a collision.

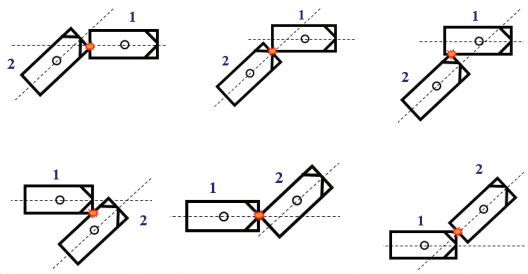


Figure 4 Possible collision types for the same approach angle (adopted from *van der Horst, 1990)*

TTC is a continuous parameter and may be calculated for any moment as long as the road users are on a collision course. It is quite widely used and some variations of this parameter have also been proposed, for example, TTCA (taking into account the acceleration of road users, *van der Horst, 1990*), inverse TTC (1/TTC, *Kiefer et al., 2005*), Time Exposed TTC and Time Integrated TTC (complex parameters taking into account the time road users spend on a collision course with TTC below a set threshold level, *Minderhoud & Bovy, 2001*).

4.2. Time Advantage

Time Advantage (further abbreviated as TAdv) is an indicator used to describe situations where two road users pass a common spatial zone, but at different times and thus avoid a collision course and thereby collision (*Hansson, 1975*). Proposed initially as a measure describing "normal" traffic conditions, Time Advantage may be seen as an extension of a safety indicator called Post-Encroachment Time (PET). The

conventional definition of PET is the time between the first road user leaving the common spatial zone and the second arriving at it (Figure 5, *Allen et al., 1977*). Thus, PET for an encounter has a single value and may be observed and measured directly. TAdv broadens the concept of PET, saying for each moment what the PET value is expected to be if the road users continue with the same speeds and paths.

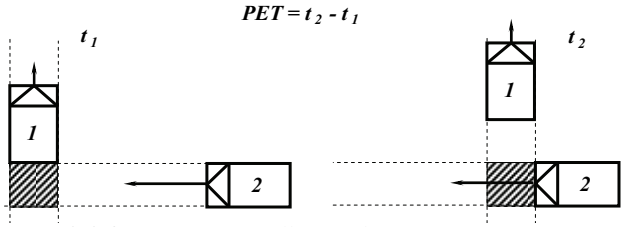


Figure 5. Conventional definition PET *(Allen et al., 1977)*

The conventional geometry-based definitions of PET and TAdv are difficult to apply when the vehicle trajectories do not cross at a right angle (which is not unusual in real life). The entrance and exit from the "common zone" are no longer time moments but periods, and it is even possible that both road users appear in the "common zone" but still avoid collision (Figure 6).

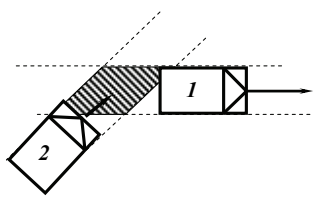


Figure 6. Problem with "geometrical" definition of PET and TAdv – both road users appear in the common zone but avoid collision.

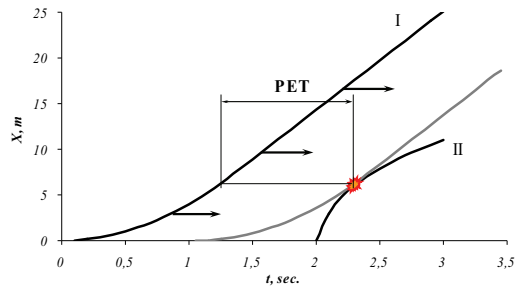


Figure 7. "Delay"-based definition of PET.

To overcome this problem other, non-geometrical terms have to be used. We propose defining PET as the *minimal delay* of the first road user which, if applied, will result in a collision course and a collision (assuming, similar to TTC, that, apart from the delay, the road users otherwise continue with the same speeds and paths). Figure 7 helps to explain this definition. Lines I and II describe the movements of two road users over the time (for simplicity we consider only one dimension and neglect the physical size of the road users). The "delay" of road user I means that its travel line has to be shifted along the time axis until it touches line II. The length of the time shift here is the Post-Encroachment Time. Time Advantage is defined in the same manner, but using the *predicted* travel lines instead. Obviously, the position of the contact point between lines I and II depends on their shapes, which in turn might

differ depending on the moment for which the prediction is made. Further, we will refer to this point as an *avoided collision point*.

In practical calculations, when the dimensions of the road users are taken into account, the TAdv has to be calculated for each possible side-corner combination. The calculation procedure for this is also provided in Appendix. For the same reason as in the case of TTC, the lowest found TAdv-value is used.

The specific of Time Advantage is that while its low values may reflect the safety aspects, the higher values (above 2-3 seconds) describe the normal traffic conditions and may be seen as a measure of one road user's power (advantage) over the other in a competition over the same spatial zone (*Hansson, 1975*). A road user having a large time advantage is most likely to be the one to pass the common zone first. However, if the time advantage is small, the second road user may accelerate with the aim of passing first instead, which occurs primarily when one of the road users is "stronger" than the other, for example, in the case of a private car vs. a pedestrian (*Várhelyi, 1998*) or a truck vs. a private car. The important point here is that the use of the same indicator to describe both safety and efficiency of the traffic processes has certain advantages and may help to better understand how these two qualities are balanced by the road users and to verify the hypotheses of such a relation (*Svensson & Hydén, 2006, Svensson, 1998, Näätänen & Summala, 1976*).

4.3. Supplementary parameter T_2 to Time Advantage

Time Advantage is by itself not sufficient to describe the collision risk since it is also important to know how soon the encroachment will occur. Even if TAdv is small at a certain moment, the road users might have plenty of time to adjust their speeds and trajectories and increase it. As an indicator describing the nearness of the encroachment, the time of the second road user arriving at the "avoided collision point" is proposed (this parameter is further abbreviated as T_2).

To use the second road user appears to be more safety-relevant as his/her arrival at the potential collision point is the very last necessary condition for a collision to occur. Whatever the actions of the first road user, it is the one who arrives last who has the largest margin, i.e., most time to take an evasive action. However, if the moment of the first road user leaving is of interest, it can be easily calculated as ($T_2 - TAdv$).

Another important property of T_2 is that it provides "smooth" transfer between the "collision course"- and "crossing course"-situations. At the moment of transfer from "collision course" to "crossing course" the TTC ceases to exist, but Time Advantage still equals zero. This makes T_2 equal to TTC, and if both TTC and T_2 are plotted on the same graph, they will make a continuous curve. Similarly to TTC, T_2 "jumps" into infinity if the second road user comes to a complete stop.

4.4. Time Gap

Depending on the relation between road users' trajectories and speeds, the "collision point" or "avoided collision point", for which TTC or TAdv are calculated, may be far ahead while the actual distance between road users might be not as large. This is

especially noticeable when the road users' trajectories are parallel or close to that (Figure 8).

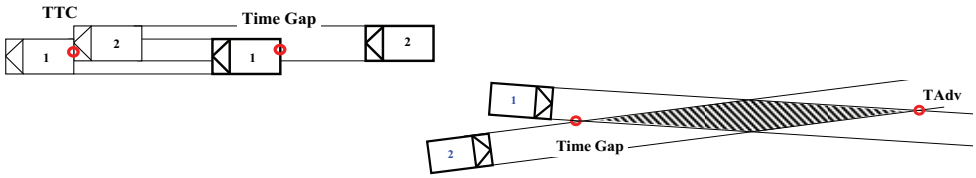


Figure 8. The points for which Time Gap, Time-to-Collision and Time Advantage are calculated.

Time Gap (TG) is a parameter that describes the actual distance between road users expressed in time units. In its conventional definition it is applied to vehicles following in a flow, and is measured as the time between the moment of the rear end of the first vehicle passing a certain point on a road and the front of the following vehicle arriving at that point (*Vogel, 2002*). This definition implies that TG is a single value measured directly at a certain location. To make it continuous and more in line with other indicators that have been discussed, a "predicted" Time Gap can be used, i.e., "the Time Gap that will be measured if the road users continue with the same speed and path".

Still, the conventional definition of TG is difficult to apply if the road users are not on exactly parallel courses. To extend this parameter and preserve its main concept, we propose the following definition. Imagine that the first of the road users is delayed to such an extent that they start moving on a collision course. There are many possible collision points, depending on the size of the delay. Of all the possible combinations of delay and proximity to collision point, the delay that produces the closest (in time) collision point is chosen. The Time Gap here is the time necessary for the second road user to arrive at the collision point. This definition includes the case of following along a parallel course, but can also be applied for any cases of overlapping courses.

As it is not known in the general case which road user is "the first" and what type of collision is the nearest in time, all possible combinations of road users' sides and corners have to be considered. The calculation procedure for TG between a point and a line section is provided in the Appendix. Again, the minimal TG value found among all the combinations is to be used.

Time Gap, presumably, has a weaker connection to collision risk compared to TTC, since it only considers the spatial proximity between the road users (in time units), but not their relative speeds. Still, it can be used for detection of potential risks at earlier stages of an encounter. This can be explained by an example of two vehicles moving on parallel courses at the same speed (i.e. no collision course exists). If the first one starts braking, the vehicles suddenly enter a collision course and the pace of the TTC decrease depends highly on the size of the time interval between the vehicles (i.e. TG). Thus, TG reflects the probability of TTC quickly reaching low values *if* the road users get onto a collision course.

4.5. Speed

Even though time-based indicators reflect both the spatial proximity and speed of the road users, one piece of important information is still missing. This can be shown by a simple example. Imagine two pairs of vehicles on a collision course, but in the first case the vehicle-speed is 10 m/s and in the second case 20 m/s. When TTC reaches 1.8 seconds in both cases, the drivers detect the risk and start braking with maximal deceleration of 6 m/s^2. In the first case they will manage to stop after 1.6 seconds and avoid collision, in the second case they will crash with a collision speed of 9 m/s.

This example clearly illustrates that time-based indicators are not sufficient to describe the severity of an encounter and need to be complemented with some speed-related indicator. The way a road user adjusts speed during a passage (road user's speed profile) also provides important behavioural information and describes the encounter as a process (*Laureshyn et al., 2009*). For these reasons we include the speed of both road users in the indicator set.

5. TWO EXAMPLES - CROSSING AND FOLLOWING COURSES

Figure 9 illustrates how the proposed indicator set describes the interaction between two road users. The first example illustrates an encounter between a car and a pedestrian on a pedestrian crossing (the pedestrian has priority). First (phase I), the car has a time advantage as the pedestrian hesitates and keeps a very low speed. Then, however, the pedestrian decides to go first and increases speed to normal pace. The TAdv of the car goes rapidly down to zero and from moment t_2 they are on a collision course (phase II). TTC is decreasing as they approach each other. Having noticed the pedestrian's behaviour, the driver brakes and from moment t_3 they are no longer on a collision course and TAdv (now the pedestrian's) starts gradually growing from zero (phase III). From moment t_4 the pedestrian is no longer in the way of the car and none of the indicators can be calculated. In this example the TG curve follows the TTC and T_2 curve and does not contribute much additional information.

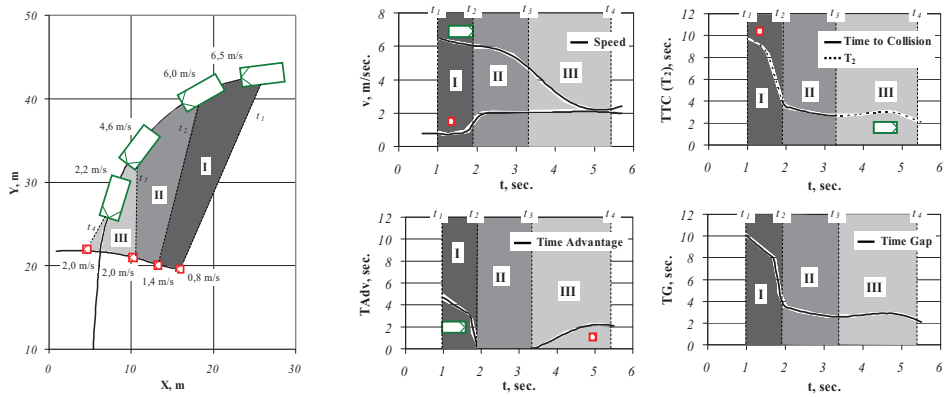
In the second example the two cars appear on a parallel course until their trajectories diverge. The speed of the following car (marked as 2 in Figure 9) is higher and the "avoided collision point" (for which the TAdv is calculated) lies in the area of trajectory divergence. Here the distance between the two cars is shorter than the distance to the "avoided collision point", and the TG curve goes lower than T_2.

6. DISCUSSION

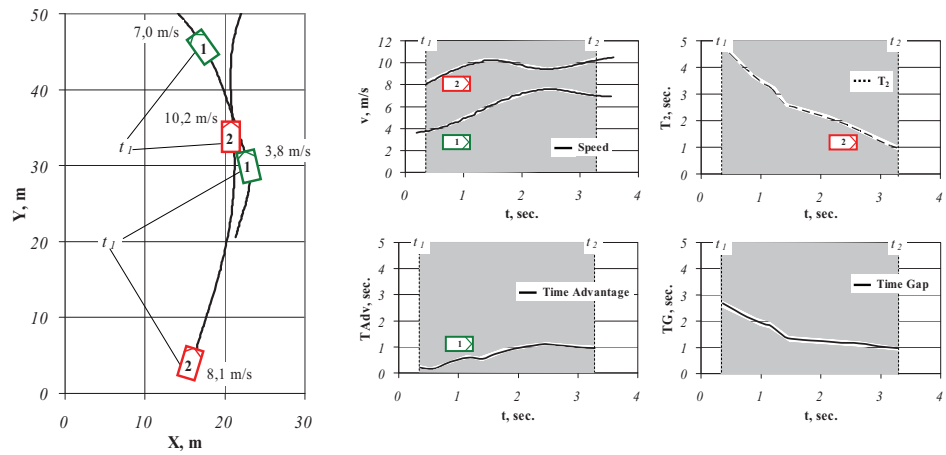
To find a universal indicator that is applicable to any type of situation during the encounter process and reflects all the relevant aspects is not simple; most probably, a set of indicators is necessary. On the other hand, it is important to keep the number of indicators as low as possible (at the risk of losing some information), otherwise it will be difficult to make the method operational.

As a first approximation, we propose a set of indicators to describe the process of an encounter between two road users. The suggested indicators address some of the

problems outlined in the theoretical framework. Many issues, however, have to be further elaborated on.



a) car-pedestrian



b) car-car (parallel course)

Figure 9. Two examples of parameter calculations for an interaction.

It is assumed here that an elementary event in traffic, which might result in an accident, is an encounter between two road users. This assumption excludes single accidents, i.e., situations with only one road user involved, even though such accidents do happen. The single accidents can be divided into several types. The first type is the accident where a second road user is involved. The first, in trying to avoid a collision with the second, drives off the road and collides with an object such as a tree. In this case the encounter does actually take place and the proposed indicators may be applied to describe it. If no other road user is present, the situation with an "active" driver and external factors that lead to an accident (e.g. an animal suddenly jumping onto the road, unexpected ice on asphalt or vehicle malfunctioning) and a "passive" driver and internal factors (alcohol influence or fatigue, resulting in loss of attention or falling asleep) may be distinguished. An approach similar to what is proposed may still be used here, i.e., the existence of a severity hierarchy for such

events may be assumed, e.g. accidents, near-accidents when the driver manages to regain the control of a vehicle on the ice or "wakes up" at the last second to avoid a crash and so on. However, the problem of integrating single road-user events and encounters into one common severity hierarchy needs to be elaborated.

There is a great variety of encounter types, with and without collision course, different approaching angles and types of road users, etc. We argue, for instance, that "diverging course" situations also have to be included in the severity hierarchy and some indicators covering these types of situations have to be developed. For a collision to become possible, road users on a diverging course need to change the relation in space (to get the paths to overlap) and, possibly, in time (to create a collision course). The proposed set of indicators allows for a continuous description and a "smooth" transfer between "collision course" and "crossing course" situations, but the "diverging course" must also be included. The new indicators have to reflect all these aspects. The challenge, however, is to unify these indicators into some general severity measure, i.e., to place them into one common severity hierarchy.

There is an advantage in using indicators that can be calculated for any moment during an encounter since the development of an encounter as a process may then be studied. On the other hand, a decision has to be taken on what moment, or combination of moments, characterises the severity of an encounter in the best way. Several options are possible. *Hayward, 1971*, and *van der Horst, 1990*, use the minimal TTC value during an encounter (TTC_{min}) to show how close to a collision road users actually come. The Swedish Conflict Technique (*Hydén, 1987*) utilises the TTC value at one point in time. It rates the severity based on the TTC value at the moment the first evasive action is started by one of the road users (called Time-to-Accident, TA). Hydén argues that TA reflects the critical moment when the road user has just detected the risk and has to take the first decision on how to act depending on the perception of severity of the event. The problem with values taken at a certain time is that they do not incorporate any information before or after the chosen moment, creating a risk that even very different encounters might be classified in the same category. Alternative non-momentarily-based measures may well be the time spent on a collision course or even more complex parameters, for example time-based indicators as proposed by *Minderhoud & Bovy, 2001*.

Another approach that may be used to classify the encounters with regard to severity is to analyse the shape of the continuous indicator profiles, i.e., TTC or TAdv curves, etc. A detailed analysis of these processes will perhaps reveal the "typical" shapes characterising the critical situations. Similarly, shapes reflecting "normal" (non-critical) processes can be found. These types of analyses can use, for example, pattern recognition methods as discussed in *Laureshyn et al., 2009*.

Many of the proposed indicators are based on predictions of a collision point in terms of the planned path and current speed of the road users. A human observer can easily project the "planned" trajectory, but it is quite difficult to explain exactly how this projection is done. A possible approximation is to assume that a road user actually follows the planned path, i.e., to use the known trajectory. This may be misleading in case the road user avoids a conflict by changing the planned path, for example taking

a larger radius in a turn or changing lanes. Another alternative is to use an "average" path, calculated from the trajectories of many road users making the same manoeuvre. The problem, however, is that in critical situations the paths might not follow the average pattern. A detailed analysis of critical situations might reveal when the deviation from the "average" pattern starts to develop during an encounter, and if the high severity of the situation can be detected *before* that moment, i.e., when the "average" assumptions are still valid. It would also be interesting to compare these objective descriptions with human observers' perceptions of the situation in order to pinpoint the relevance of suddenness and closeness to collision point when diverging from the "average" path.

van der Horst, 1990, tests different variations of TTC definitions based on assumptions of constant angular velocity and constant acceleration of a vehicle (this is supposed to represent a situation when a driver is no longer controlling the vehicle and the steering wheel and the gas pedal positions are kept unchanged). The paths calculated with constant angular velocity easily take very peculiar shapes and lead outside the road. As for constant accelerations, the TTC values are still reasonable, but there is no clear evidence to show that the predictive power of TTC improves.

The methods for combining the accident risk and the severity of consequences into one severity measure are still missing. In the Swedish Conflict Technique (*Hydén, 1987*) this problem is circumvented by defining the severity of an encounter as a potential of an *injury* accident, i.e. property-damage accidents are not considered at all. However, there are indications of the existence of a more universal severity measure. During the calibration study of traffic conflict techniques from different countries (*Grayson, 1984*), the severity rating of the conflicts based on objective measures was compared with subjective ratings of the human observers. A stronger agreement was found among the subjective ratings of different observers than among the objective ratings based on definitions of the techniques tested. One of the explanations offered was that human observers considered both the collision risk and consequences, while the objective measures often reflect just one of the aspects (in most cases, the collision risk). The challenge, however, is to find an objective and operational measure that corresponds to the subjective severity judgements. This will probably lead to more valid conflict measures.

Our hypothesis is that a feasible way forward regarding the description of the severity of the encounter process is to elaborate further on the shape of the severity hierarchy and the assumption of a "true" hierarchy. Nonetheless, the set of indicators describing the severity of the process has to be translated into one common severity measure for inclusion in one common severity hierarchy. It is important to further elaborate on what the whole shape and the accumulation of events at different levels represent. The severity hierarchies proposed earlier (*Svensson & Hydén, 2006, Svensson, 1998*) were only based on events with collision courses. It was argued, for instance, that interactions with fairly high severities could be positive from a safety point of view because they were frequent and severe enough to increase awareness, and that these events were predominant at a site without serious conflicts and accidents. It will be interesting to analyse whether the same interpretation is valid for the hierarchies proposed here. With information about the encounter processes and the severity of

these processes it will be possible to formulate and test hypotheses on the interrelationships of design of the traffic environment, behaviour and risk. It is also important to point out that the behaviour described by the severity hierarchies could reflect other qualities in traffic besides safety, like mobility and the desire to balance these qualities

Our expectation is that video analysis is the tool that will provide the necessary micro-level data on the behaviour of road users. With automated video analysis it will be possible to elaborate on the severity hierarchies as long-term recording will provide us with accidents as well. However, some important indicators that have direct implications for the accident risk, and especially the severity of the accidents (e.g. use of helmets, road user age, eye contact and other signals sent by road users) are hard to extract from video data (*Laureshyn & Svensson, 2009*). It might be necessary to complement video analysis with some other data collection method (e.g. human observers who look through a video that has been initially automatically filtered) to get the necessary information.

When introducing a new method and new indicators, the most important aspects are their reliability and validity (*Laureshyn & Svensson, 2009*). The reliability is a property to produce the results of the same accuracy irrespective of where and in what conditions and by whom measurements are made, thus ensuring that the difference in the results is attributable to the difference in the studied quality (safety) and not to a measurement error. The validity guarantees the robust relation of the used indicators with the studied quality. Establishing these two qualities is to be seen as a necessary step in the development of the proposed method. Again, with automated video analysis and a framework around relevant indicators, it will be feasible to elaborate on validity and reliability.

7. CONCLUSIONS

An encounter between road users is a process that can be described as a continuous interplay over time and space. For some encounters the road users are on a collision course, while in other encounters there is no collision course. As a further complication, moving on/ not moving on a collision course may change during the encounter process. Most prevalent traffic safety indicators do not consider the severity of the whole process, but assign a severity to a certain moment during this process without considering occurrences just before or after this moment. Moreover, safety is far too often treated in a one-dimensional manner as if it is the only motive while moving in traffic. Hence, other motives like efficiency and comfort are not considered. This paper elaborates on how to improve the calculation of the existing indicators like Time-to-Collision, Time Gap and Time Advantage for all types of approach angles and complement them with speed and T_2 in order to estimate the severity. The paper also suggests how to make smooth transfers between the different indicators in order to reveal the whole encounter process as a continuous interplay between road users. By describing the processes preceding accidents with regard to TTC, Time Gap and Time Advantage, speed and T_2, as well as "normal" encounters not resulting in accidents or serious conflicts, it will be possible to analyse what

distinguishes a "safe" process from an "unsafe" process. With this information it will be possible to organise all encounters in a severity hierarchy and classify the severity of the encounters with regard to the whole encounter process. This will be a considerable contribution to increasing knowledge of the traffic safety process and understanding road users' trade offs between safety and efficiency in traffic.

APPENDIX

Calculation of Time-to-Collision for a point and a line section

Let (x_p, y_p) be the current co-ordinates of the point and v_p its speed vector (Figure 10). Then the position of the point at an instant time t can be described by equations:

$$\begin{cases} x_p = x'_p + v_{px} \cdot t \\ y_p = y'_p + v_{py} \cdot t \end{cases}, \qquad \textit{Eq. 4}$$

where (x'_p, y'_p) is the initial position of the point and v_{px} and v_{py} are the projections of the speed vector on the X- and Y-axes.

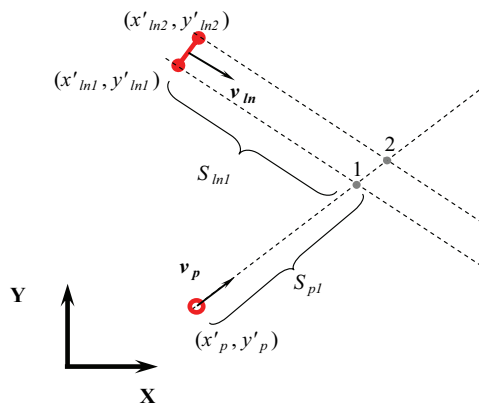


Figure 10. Calculation of TTC, TAdv and TG for a point and a line section.

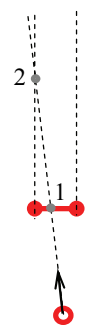


Figure 11. An example of a special case, t_{ln1} = 0.

The position of the line section ends (x_{ln1}, y_{ln1}) and (x_{ln2}, y_{ln2}) is described by equations:

$$\begin{cases} x_{\ln 1} = x'_{\ln 1} + v_{\ln x} \cdot t \\ y_{\ln 1} = y'_{\ln 1} + v_{\ln y} \cdot t \end{cases} \quad \text{and} \quad \begin{cases} x_{\ln 2} = x'_{\ln 2} + v_{\ln x} \cdot t \\ y_{\ln 2} = y'_{\ln 2} + v_{\ln y} \cdot t \end{cases}, \qquad \textit{Eq. 5}$$

where (x'_{ln1}, y'_{ln1}) and (x'_{ln2}, y'_{ln2}) are the initial positions of the line section ends; $v_{ln\,x}$ and $v_{ln\,y}$ are the projections of the line speed vector v_{ln} on the X- and Y-axes.
The line equation in its canonical form is

$$x - x_{\ln 1} = \frac{y - y_{\ln 1}}{k}, \qquad \textit{Eq. 6}$$

where k is a parameter describing the line slope. In case where the line moves parallel to itself *k* remains constant and thus can be found from the initial line position as

$$k = \frac{y'_{\ln 2} - y'_{\ln 1}}{x'_{\ln 2} - x'_{\ln 1}}; \qquad \textit{Eq. 7}$$

Substituting the point co-ordinates into the line equation, the time of the collision can be found as:

$$t_{coll.} = -\frac{(y'_p - y'_{\ln 1}) - k \cdot (x'_p - x'_{\ln 1})}{(v_{py} - v_{\ln y}) - k \cdot (v_{px} - v_{\ln x})}, \qquad \textit{Eq. 8}$$

or, in case where the denominator in Eq. 7 is zero ($k \to \infty$):

$$t_{coll.} = -\frac{x'_p - x'_{\ln 1}}{v_{px} - v_{\ln x}}. \qquad \textit{Eq. 9}$$

Here only positive $t_{coll.}$-values are of interest. The condition for the point to cross the line within the section is that at the moment $t = t_{coll.}$:

$$\begin{cases} x_{\ln 1} \le x_p \le x_{\ln 2} \ \ if \ \ x_{\ln 2} \ge x_{\ln 1} \\ x_{\ln 2} \le x_p \le x_{\ln 1} \ \ if \ \ x_{\ln 1} \ge x_2 \end{cases} \text{ and } \begin{cases} y_{\ln 1} \le y_p \le y_{\ln 2} \ \ if \ \ y_{\ln 2} \ge y_{\ln 1} \\ y_{\ln 2} \le y_p \le y_{\ln 1} \ \ if \ \ y_{\ln 1} \ge y_2 \end{cases}. \qquad \textit{Eq. 10}$$

The collision point ($x_{coll.}$, $y_{coll.}$) coincides with the point position (x_p, y_p) at the moment $t = t_{coll.}$.

Calculation of Time Advantage and Time Gap for a line section and a point

To find the time distance between a line section and a point, it is enough to check the time differences with which the point and the line section ends pass the common points 1 and 2 (Figure 11). For point 1 the time difference Δt_1 is:

$$\Delta t_1 = \left| t_{p1} - t_{\ln 1} \right| = \left| S_{p1} / v_p - S_{\ln 1} / v_{\ln} \right| , \qquad \textit{Eq. 11}$$

where t_{p1} and t_{ln1} are times necessary for the point and the section end to reach the common point 1;

S_{p1} and S_{ln1} are the distances from the point and the line section end to the common point 1.

Calculations of the time difference Δt_2 for point 2 are done in the same way. Time Advantage will be the minimal value between Δt_1 and Δt_2.

A special case when the point and a section end trajectory do not cross is shown in Figure 11. In this example $t_{ln1} = 0$ and $\Delta t_1 = t_{p1}$. Similarly, the cases when t_{p1}, t_{p2} and t_{ln2} are equal to zero have to be considered.

To find Time Gap, points 1 and 2 have to be checked first. The "second" road user takes a longer time to arrive at a common point. For example, the time necessary for the "second" road user to arrive at point 1 is

$$t_{2-1} = \max\left(t_{p1}; t_{\ln 1}\right). \qquad \textit{Eq. 12}$$

The time $t_{2\text{-}2}$ necessary for the "second" road user to arrive at point 2 is calculated in the same way. Time Gap will be the minimal value between $t_{2\text{-}1}$ and $t_{2\text{-}2}$. The special cases such as the one shown in Figure 11 also have to be considered.

These calculations of TG are performed in cases where the point and the line section are not on a collision course. If TTC can be calculated, the Time Gap is equal to TTC.

REFERENCES

Allen, B. L., B. T. Shin, P. J. Cooper (1977). Analysis of traffic conflicts and collisions. Department of Civil Engineering, McMaster University.

Asmussen, E. (Ed.) (1984). International calibration study of traffic conflict techniques. Springer-Verlag, Berlin Heidelberg New York Tokyo.

Carlsson, G. (2004). Kunskapssammanställning - dödsolyckor och hastighet (*in Swedish*) Knowledge summary - fatal accidents and speed. NTF-kansliet.

Englund, A., N. P. Gregersen, C. Hydén, P. Lövsund, L. Åberg (1998). Trafiksäkerhet – en kunskapsöversikt (*in Swedish*) Traffic safety - a knowledge review. KFB, Studentlitteratur, Lund.

Gettman, D., L. Head (2003). Surrogate safety measures from traffic simulation models. U. S. Department of Transportation, Federal Highway Administration. Report No. FHWA-RD-03-050.

Grayson, G. B. (1984). The Malmö study: a calibration of traffic conflict techniques. SWOV. Report R-84-12.

Hansson, A. (1975). Studies in driver behaviour, with applications in traffic design and planning. Two examples (*in Swedish*). Lund Institute of Technology, Department of Traffic Planning and Engineering, University of Lund. Bulletin 11.

Hayward, J. C. (1971). Near misses as a measure of safety at urban intersections. Doctoral thesis, The Pensilvania State University, Department of Civil Engineering.

Hydén, C. (1987). The development of a method for traffic safety evaluation: the Swedish traffic conflict technique. Doctoral thesis, Lund University, Department of Traffic Planning and Engineering.

Hydén, C. (1996). Traffic conflicts technique: state-of-the-art *in* H. H. Topp (Ed.) Traffic safety work with video-processing. Transportation Department, University Kaiserlauten.

Kiefer, R. J., D. J. LeBlanc, C. A. Flannagan (2005). Developing an inverse time-to-collision crash alert timing approach based on drivers' last-second braking and steering judgments. Accident Analysis and Prevention 37, pp. 295-303.

Kraay, J. H., A. R. A. van der Horst (1985). The Trautenfels study: a diagnosis of road safety using Dutch conflict observation technique DOCTOR. Institute for Road Safety Research SWOV, The Netherlands. R-85-53.

Laureshyn, A., H. Ardö, T. Jonsson, Å. Svensson (2009). Application of automated video analysis for behavioural studies: concept and experience. IET Intelligent Transport Systems 3 (3), pp. 345-357.

Laureshyn, A., Å. Svensson (2009). Road user behaviour indicators in automated video analysis systems. Journal of ITS (***submitted***).

Laureshyn, A., K. Åström, K. Brundell-Freij (2009). From speed profile data to analysis of behaviour: classification by pattern recognition techniques. IATSS Research 33 (2), pp. 88-98.

Messelodi, S., C. M. Modena, M. Zanin (2004). A computer vision system for the detection of vehicles at urbam road intersections. Instituto Trentino di Cultura (ITC). Technical Report T04-02-07.

Minderhoud, M. M., P. H. L. Bovy (2001). Extended time-to-collision measures for road traffic safety assessment. Accident Analysis and Prevention 33, pp. 89-97.

Näätänen, R., H. Summala (1976). Road-user behavior and traffic accidents. North-Holland Publishing Co, Amsterdam.

Svensson, Å. (1998). A method for analysing the traffic process in a safety perspective. Doctoral thesis, University of Lund, Lund Institute of Technology, Department of Traffic Planning and Engineering.

Svensson, Å., C. Hydén (2006). Estimating the severity of safety related behaviour. Accident Analysis & Prevention 38, pp. 379-385.

van der Horst, R. (1990). A time-based analysis of road user behaviour in normal and critical encounters. Doctoral thesis, Delft University of Technology.

van der Horst, R., R. J. Kraay (1986). The Dutch conflict technique – DOCTOR. ICTCT Workshop, Budapest.
Várhelyi, A. (1998). Driver's speed behaviour at a zebra crossing: a case study. Accident Analysis and Prevention 30 (6), pp. 731-743.
Wilde, G. J. S. (1994). Target Risk. PDE Publications, Toronto.
Vogel, K. (2002). What characterizes a "free vehicle" in an urban area? Transportation Research Part F 5, pp. 15-29.

Paper III

(2009) IATSS Research 33 (2), pp. 88-98

FROM SPEED PROFILE DATA TO ANALYSIS OF BEHAVIOUR: CLASSIFICATION BY PATTERN RECOGNITION TECHNIQUES

Aliaksei Laureshyn[a], Kalle Åström[b], Karin Brundell-Freij[c]

[a] Traffic and Roads, Deptartment of Technology and Society, Faculty of Engineering, LTH, Lund University, Sweden

[b] Centre for Mathematical Sciences, Faculty of Engineering, LTH, Lund University, Sweden

[c] WSP Group Sweden, Analysis & Strategy

Abstract: Classification of speed profiles is necessary to allow interpretation of automatic speed measurements in terms of road user behaviour. Aggregation without considering variation in individual profile shapes easily leads to aggregation bias, while classification based on exogeneous criteria runs the risk of loosing important information on behavioural (co-)variation. In this paper we test how three pattern recognition techniques (cluster analysis, supervised learning and dimension reduction) can be applied to automatically classify the shapes of speed profiles of individual vehicles into interpretable types, with a minimum of *a priori* assumptions. The data for the tests is obtained from an automated video analysis system and the results of automated classification are compared to the classification by a human observer done from the video. Normalisation of the speed profiles to a constant number of data points with the same spatial reference allows them to be treated as multidimensional vectors. The *k*-means clustering algorithm groups the vectors (profiles) based on their proximity in multidimensional space. The results are satisfactory, but still the least successful among the tested techniques. Supervised learning (nearest neighbour algorithm tested) uses a training dataset produced beforehand to assign a profile to a specific group. Manual selection of the profiles for the training dataset allows better control of the output results and the classification results are the most successful in the tests. Dimension reduction techniques decrease the amount of data representing each profile by extracting the most typical "features", which allows for better data visualisation and simplifies the classification procedures afterwards. The singular value decomposition (SVD) used in the test performs quite satisfactorily. The general conclusion is that pattern recognition techniques perform well in automated classification of speed profiles compared to classification by a human observer. However, there are no given rules on which technique will perform best.

Keywords: Speed profile, behaviour analysis, pattern recognition, clustering, supervised learning, dimension reduction.

1. INTRODUCTION

The main aim of studying the behaviour in traffic is to understand the reasons for road users to act as they do and learn factors that can affect their actions (*Englund et*

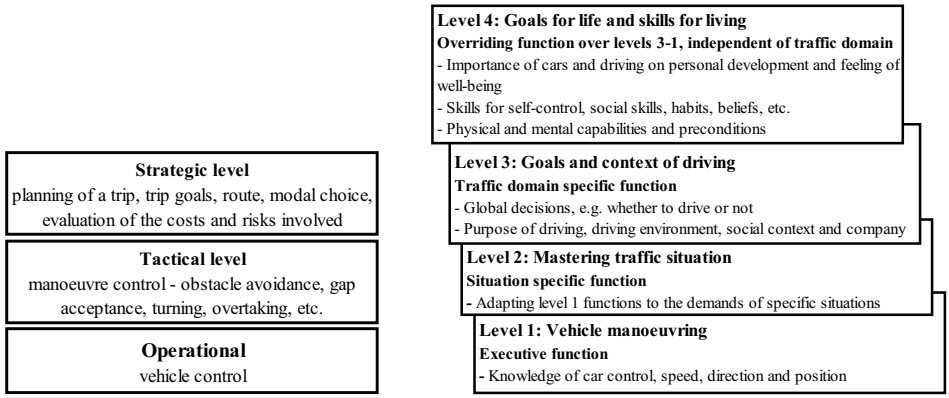
al., 1998). One of the main behaviour characteristics is the choice of speed, which has direct implication on safety, efficiency and the size of environmental impacts both for an individual road user and the traffic system in general (*Aarts & van Schagen, 2006, Nilsson, 2004, Ericsson, 2000, Hagring, 2000, Kloeden et al., 1997*). Not only the general speed level, but also the way the speed changes is of importance. For example, intensive accelerations of a vehicle cause much higher emissions, while braking and the suddenness of braking – jerk – are often regarded as indicators of conflict situations (*Larsson, 2009, Malkhamah et al., 2005, Nygård, 1999*). Investigation of the relations between the speed and design of roads and intersections can give an idea of how the optimal speed regime can be reached (*Karlgren, 2001, Pau & Angius, 2001, Hydén & Várhelyi, 2000*).

From the psychological perspective, the speed we observe in any given moment can be seen as a result of the many tasks a road user performs. These can be regarded to be related to decision-making on different levels. For example, *Michon, 1985*, describes the following model. The first level (see Figure 1*a*) is operational and relates to control of the vehicle and decisions about the use of the steering wheel and pedals, gear choice, etc. The second, tactical, level refers to manoeuvring and immediate interactions with other road users. The third, upper level in the hierarchy is strategic and it concerns tasks like trip planning, navigation and route choice. A more recent model, the GADGET-matrix (named after the European project for which it was developed, see Figure 1*b*), suggests also the fourth level, described as "goals for life and skills for living" and referring to social skills, beliefs, importance of driving for personal well-being and social status, etc. (*Peräaho et al., 2003*).

In the view of the described cognitive models, the momentary speed of a road user reflects performance on the lowest operational level. The behaviour, i.e. road user actions in relation to other road users and the road, belongs rather to the second, tactical level and results in a continuous process of speed adaptation to the traffic environment. To understand the behaviour it is necessary to relate the speed changes to the traffic conditions. One of the challenges, however, is to extract the behavioural information from the extensive operational speed data.

One way of collecting the behavioural data is to use qualitative description of the speed changes. An observer makes a note, for example, that a road user "slows down", "stops", "does not alter speed", "yields to another road user", "drives on red", etc. (*Martinez & Porter, 2006, Hakkert et al., 2002, Carsten et al., 1998*). The use of human observers put serious limitations on amount of data that can be practically collected and also on how detailed the data could be. On the other hand, humans judge the observed situations holistically, and in classifying them might consider dimensions not even captured by the objective measurements of speed or other variables. A typical example of judging traffic situations is traffic conflict studies, where observers are found to be very good in distinguishing between serious and non-serious conflicts and also rating the conflicts' severity. In fact, in an international calibration study there was found a higher agreement between individual observers' judgements than between observers and various objective measures of conflict severity. This only supports the hypothesis that there are relevant subjective

dimensions that play a role in the interpretation of observed processes in traffic (*Hydén, 1987, Grayson, 1984*).



a) *b)*

Figure 1. The hierarchy of the road user's tasks: a) adapted from** Michon, 1985**; b) Gadget-matrix** (Peräaho et al., 2003)**.

Another tradition in studying speed data is to use speed loggers installed in vehicles driving in real traffic. After identification of the important characteristics, the data is interpreted in terms of driving patterns and used, for example, to produce standard driving cycles for vehicle tests and estimation of the emission factors (*Larsson, 2009, Ericsson, 2000, André et al., 1995*). The problem with this method is that the studied population of drivers and vehicles is normally very limited and the information about the traffic situation (e.g. the presence of other road users) is missing.

In recent time, video analysis techniques are becoming a popular tool for traffic data collection (*Laureshyn et al., 2009, Parkhurst, 2006, Messelodi et al., 2004*). This technology provides an opportunity to measure speed with high time frequency and for large populations of road users. However, the problem of interpretation in behaviour-terms still remains, and the greatly increased amount of data that is collected necessitates the analysis to be automated.

Simple aggregation to an average (or 85-percentile) speed profile (*Karlgren, 2001, Várhelyi, 1998*) loses information on differences between the individuals and correlation (over time) within individual profiles. It may therefore lead to aggregation bias and misleading final profile shape if the individual profiles are very different in character. *Sekine & Sekine, 2009*, propose LUNA (Location UNiversal Archive-format) aggregation format, which provides speed distributions at several points along the studied section. This preserves to a greater extent the variation of speeds at each cross-section point, but the longitudinal connections between the points of individual profiles is still lost. None of these approaches utilises the information about the variation in shapes of individual profiles, which can be attributed to different behavioural strategies. If profiles are classified before the aggregation, classification based on exogenous (pre-set) criteria runs a risk of loosing important information on behavioural co-variation.

In order to utilise the advantages of the detailed data contained by large samples of speed profiles, it is necessary to have a method that:

I. differentiates between the behaviour types based on endogenous (derived from data) criteria in a way similar to a human observer;

II. makes use of the systematic variation in the data that can be attributed to different types of behaviour (i.e. analyses shapes of the speed profiles);

III. can handle large amounts of data as produced, for example, by the video analysis techniques.

We suggest using pattern recognition techniques to fill this methodological gap. Pattern recognition is a topic in machine learning theory that aims at classifying data based on *a priori* knowledge or information extracted from the data itself. In this paper we test three techniques (cluster analysis, supervised learning and dimension reduction) according to the criteria I - III described above.

The paper has the following structure. First, we describe the dataset that was used in the tests. Then, we shortly explain the principles of the three pattern recognition techniques one by one and apply them for classification of the speed profiles in the dataset. Finally, the performance of the techniques in the tests is discussed and the conclusions are drawn.

2. THE DATASET: LEFT TURNING VEHICLES AT A SIGNALISED INTERSECTION

The data for the tests was collected as a part of the work on developing a system for automated video analysis (*Ardö, 2009, Laureshyn et al., 2009*). A signalised intersection was video-filmed and the system was used to extract the trajectories and speeds of moving road users. Only data on left-turning cars was saved (Figure 2). The reason for limiting the vehicle type to cars only is that for larger vehicles the accuracy of speed estimated from video is not very accurate. The left-turning manoeuvre was chosen as it requires interactions with two conflicting flows (on-coming traffic and pedestrians at the pedestrian crossing) and more variation in profile shapes is found compared to other manoeuvres. The speed data was visually checked for consistency and manually corrected in cases of obvious errors in detection. After removing the incomplete profiles (for example if the tracked vehicle was occluded by other vehicles for some time), 253 profiles were left for analysis. The profiles were trimmed and adjusted so that each profile contained 60 data points, evenly distributed along the trajectory between the defined start and end lines. This allowed direct point comparison of the profiles by referring to their order numbers only.

According to the rules, a left-turning vehicle must yield both to the traffic coming from the opposite direction and to pedestrians who have green in the same phase. This results in four possible types of situations:

a) There are vehicles coming from the opposite direction, the driver has to yield by braking near the middle line.

b) There is a pedestrian at the pedestrian crossing, the driver has to brake before the crossing.

c) No conflicting traffic is present or the gap is sufficiently large, a turning vehicle proceeds with nearly constant or slightly increasing speed.

d) A driver has to brake both near the middle line and near the pedestrian crossing. This situations are extremely rare, since the pedestrian flow is low and those who are present usually manage to complete their passage while the driver is waiting at the middle line. There were no such situations in the studied dataset.

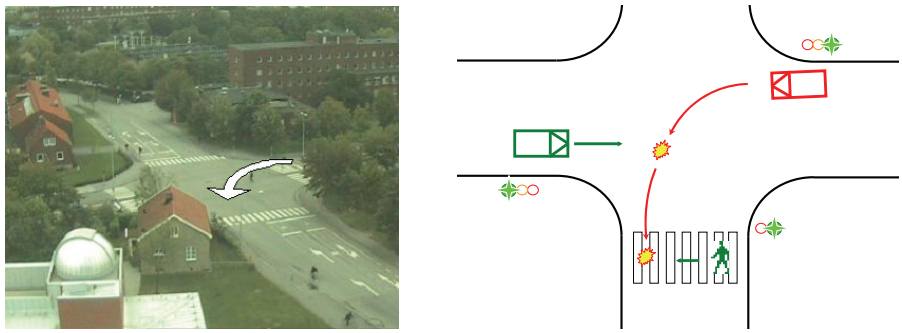
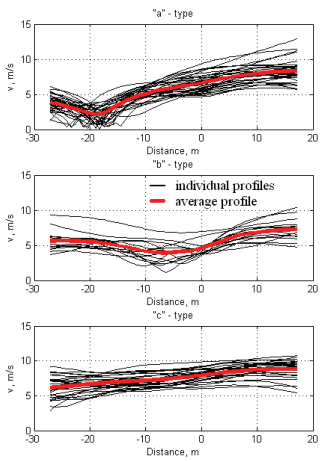


Figure 2. View of the observation site and conflict points for left turning vehicles.

The situations in the dataset were classified into three groups (*a*, *b* and *c*- types) by an observer who watched through the video clips from which the speed profiles had been extracted. To avoid influence on the observer's judgements, he was not allowed to see the speed profiles corresponding to the situations. The results of the classification are shown on Figure 3.



a)

Situation type	Number of situations	Percentage
"a"	*102*	*40%*
"b"	*16*	*6%*
"c"	*135*	*54%*
Total	**253**	**100%**

b)

Figure 3. Speed profiles in the situations classified by an observer** (a) **and the distribution of the situations by type** (b)**.

Examination of the speed profiles reveals that each type of situations has quite distinctive shape, shown on Figure 4. However, not all the profiles fit the typical

shapes perfectly and there is a large group of profiles that appear to be somewhere "in between" two shapes. The observer also expressed difficulties with classifying such situations as the description of more then one type matched them (for example, a car moves slowly forward, avoiding thus sharp baking at the middle line, but still being affected by the on-coming traffic). This is a general problem of the diversity of behaviour forms that complicates its classification. We assume here that the observer's classification is the best possible to achieve and it is used as the "ground truth" in the following tests.

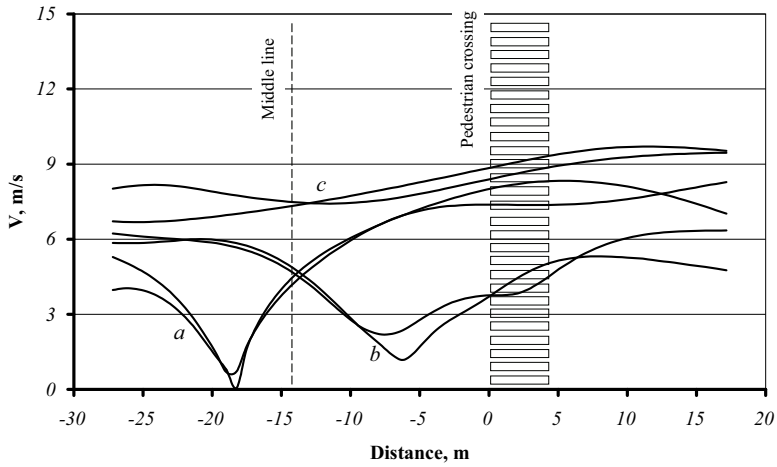


Figure 4. Three most typical profile shapes: a – no on-coming traffic or pedestrians; b – driver yields to the on-coming vehicles; c – driver yields to pedestrians at the pedestrian crossing.

3. PATTERN RECOGNITION TECHNIQUES FOR SPEED PROFILES CLASSIFICATION

3.1 Speed profile as a vector

The concept of the vector, which is defined as a geometrical object having both length and direction, is a key element in linear algebra. A vector in n-dimensional space is described by n co-ordinates, which is an ordered list of n numbers.

All the profiles in the dataset have the same number of points (60, or n in the general case) and each point i refers to the same location, i.e. the points between the profiles can be directly compared. A possible interpretation of this data is that each profile represents a vector in n-dimensional space. The speed at a point i is thus the value of the ith-co-ordinate.

This approach helps to better understand the logic of the techniques described in the following sections.

3.2 Cluster analysis

Cluster analysis is a general name for methods of dividing the data into several partitions (clusters) according to some properties considered common for the items within the cluster. Most often this property is proximity, i.e. the items in a cluster are closer to each other or to the cluster centre than to other items or other cluster centres. A clustering algorithm may force the data into a pre-defined number of clusters *k* (*k*-clustering) or find the optimal number of clusters based on the data.

The simplest clustering algorithm is a *k*-means clustering (*Duda & Hart, 1973*). Figure 5 illustrates its principle in a 2-dimensional example. First, some points are randomly chosen as cluster centres (C_1' and C_2') and each point is assigned to the nearest centre (iteration 1). Then the cluster centres are recalculated as the average of the assigned points and the same procedure is repeated (iteration 2). This is continued until a new reassignment does not differ from the previous one or until the centre's co-ordinates do not change significantly after more iterations. In case of *n* co-ordinates, the algorithm works in the same way, but the distances are calculated according to the rules of *n*-dimensional space.

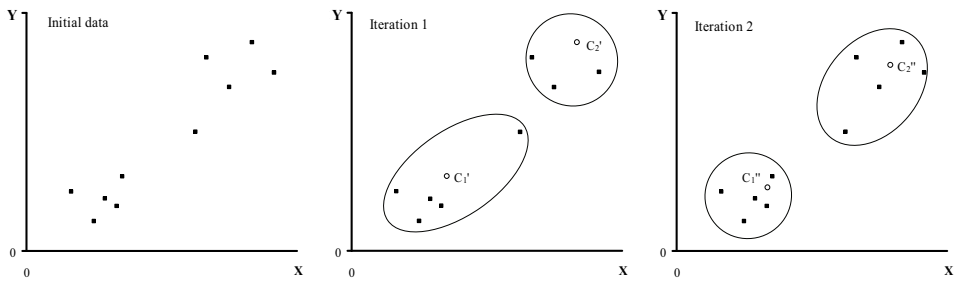
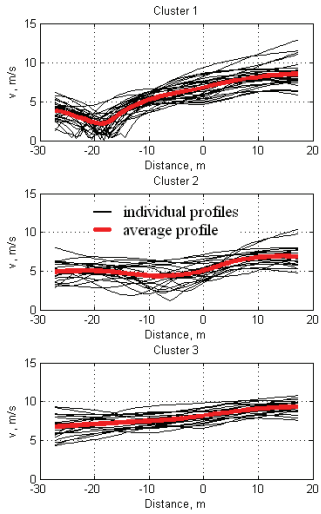


Figure 5. The k-means clustering algorithm at work.

The advantages of *k*-means are its extreme simplicity and speed of calculation, but the drawbacks are dependence of the results on the choice of initial centres, assumption of the cluster round shapes, etc. More advanced clustering algorithms, such as quality threshold (QT), expectation maximisation (EM) and hierarchical clustering treat many of these problems better, but may incur other problems (*Ripley, 1996*).

Figure 6 illustrates the results of *k*-means clustering of the speed profiles data (to keep diagrams readable only 75 profiles are plotted). Cluster 1 represents mostly the *a*-type profiles. Cluster 2 is a mixture of *b*- and *c*-types and Cluster 3 contains clearly *c*-type profiles. The result is not optimal, and there are several possible ways to improve it. First, a larger number of clusters (for example 9 instead of 3) may be used to make possible differences between the profiles more visible and then merge the clusters belonging to the same type. Another alternative is to repeat cluster analysis within the clusters (in this case it is relevant for Cluster 2), to split the mixed profiles of different types. Comparing clusters 2 and 3 one may note that the profiles in cluster 2 generally have lower speed levels than in Cluster 3. This is not surprising since the *k*-means algorithm clusters the profiles based on the distance between the points, i.e. profiles with different shapes close to each other have a higher chance of appearing in

the same cluster than profiles having the same shape but different speed levels. This may be tackled by, for example, using the derivative of speed as a co-ordinate at each point instead of speed itself, since it will produce more level difference between the profiles of different shapes.



a)

			k-means situation type			Total
			"a"	"b"	"c"	
Observer situation type		"a"	32%	8%	0%	40%
		"b"	0%	6%	0%	6%
		"c"	2%	9%	43%	54%
Total			34%	23%	43%	100%

Correct classifications - *81%*

b)

Figure 6. Results of k-means clustering of the speed profiles** (a) **and comparison with the observer's results** (b)**.

3.3 Supervised learning

Supervised learning is another technique that may be used for data classification. The main difference from clustering is, however, that the classification function is learnt from a training dataset containing both the input objects and the desired outputs. There is a wide range of classification algorithms developed, where *k*-nearest neighbours is one of the simplest. Figure 7 illustrates how the algorithm works in a two-dimensional example. The training dataset contains objects of two types – black squares and white circles. In order to decide to which type a new object belongs, the distances to all the objects in the training dataset are calculated and *k* nearest of them are selected. The decision is made based on simple voting, i.e. if the majority of the selected neighbours are squares, the new object will be classified as a square and vice versa. It is reasonable to select *k* equal to some odd number to avoid the situation of equal votes.

The algorithm described above has many drawbacks. First, the choice of *k* is crucial for the classification results. In the example, if only three neighbours are considered ($k = 3$), the new object is classified as a square, but if $k = 5$, it is classified as a circle. When the number of possible classes is more than two, situations of equal votes become possible and there is a need for additional criteria to make a decision. If a certain type of object is dominating in the training set, it may also tend to "win" the votes more often just because of the denser population and higher probability of appearing near the tested object. Other algorithms, such as support vector machines

(SVMs) and artificial neural networks (ANNs) often treat these problems better (*Ripley, 1996, Duda & Hart, 1973*).

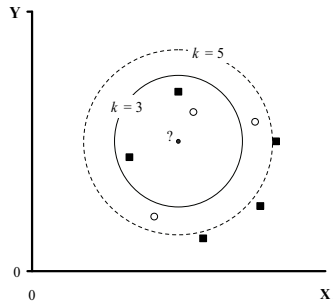
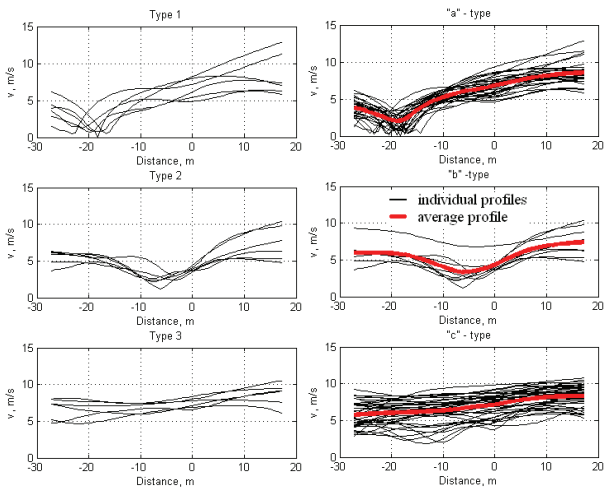


Figure 7. k-nearest neighbours classification.

Figure 8 illustrates the classification of the speed profiles using the nearest neighbour algorithm (k = 1). Six profiles of each type with very typical shapes are selected as a training dataset (Figure 8*a*) and the results of the classification are presented in Figure 8*b* (only a 75 profiles are plotted to keep the diagram readable). The results appear to be more robust compared to unsupervised clustering, but preparation of the training set requires additional manual work.



			nearest neighbour situation type "a"	"b"	"c"	Total
Observer situation type		"a"	36%	0%	4%	40%
		"b"	0%	3%	3%	6%
		"c"	2%	0%	52%	54%
Total			38%	3%	59%	100%

Correct classifications - 91%

a) *b)* *c)*

Figure 8. Training set (a), the nearest neighbour classification based on the training set (b) and comparison with the observer's results (c).

3.4 Dimension reduction

The n-co-ordinate set gives the exact description of the vector (profile shape). However, operating all n co-ordinates is not always convenient as it complicates data visualisation and analysis, increases computation time and may even yield less accurate results compared to analysis of simplified data. A possible solution is to find

an approximation to a vector that may be described by fewer co-ordinates but still preserves enough information about the vector's important features.

This problem may be solved in several possible ways. One is to select several co-ordinates from the original set, given that they contain the most important information (feature selection). Another way is to create a completely new co-ordinate system with fewer dimensions than n and project the original vector in the new space (feature extraction). The task here is to find a system that preserves as much information about the vector's features as possible and omits less important information.

To explain in a simple way how this works, let us consider a vector $\boldsymbol{a}$ in 3-dimensional space (Figure 9). Let $\boldsymbol{e_1}$, $\boldsymbol{e_2}$, $\boldsymbol{e_3}$ be an orthonormal co-ordinate system. A complete vector description is given by its three co-ordinates, i.e. its projections on the base vectors ($\boldsymbol{a} = c_1 \cdot \boldsymbol{e_1} + c_2 \cdot \boldsymbol{e_2} + c_3 \cdot \boldsymbol{e_3}$). Omitting one of the co-ordinates, for example c_3, we approximate the original vector with its projection $\boldsymbol{a_{12}}$ on the plane defined by the base vectors $\boldsymbol{e_1}$ and $\boldsymbol{e_2}$. If we omit two co-ordinates, c_2 and c_3, the original vector is approximated with its projection $\boldsymbol{a_1}$ on the base vector $\boldsymbol{e_1}$. Obviously, as more co-ordinates are omitted, the quality of the approximation decreases. However, how much information is lost with a co-ordinate depends on the chosen co-ordinate system and which of the co-ordinates are omitted. For example, if the orientation of the vector is close to one of the base vectors, this co-ordinate will be most important. Omitting it will imply that nearly all the information about the vector is lost. Other co-ordinates, on the contrary, may be easily omitted without introducing any substantial errors in vector description.

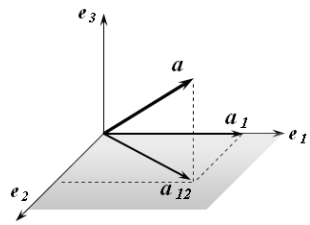


Figure 9. Approximations of a vector in 3-dimensional space.

One of the feature extraction techniques is singular value decomposition (SVD, *Strang, 1986*). Let us construct a matrix M in which each m column represents a vector of length n. According to the SVD theorem a $n \times m$ matrix M can be presented in the form of a product of three components:

$$M = U \cdot S \cdot V^T ,$$

where U and V are unitary matrices of $n \times n$ and $m \times m$ size respectively and S is a $n \times m$ diagonal matrix with non-negative values on the diagonal arranged in a non-increasing order.

One of the interpretations of the SVD results is that U is a set of orthonormal vectors defining a new co-ordinate system, while $S \bullet V^T$ are the co-ordinates of the original vectors in the new co-ordinate system. An important property of SVD is that the new base vectors are sorted in decreasing order of their importance, i.e. the first co-

ordinate gives more information about the original vectors than the second one and so on. Another important property is that if one wants to use only i co-ordinates to describe the original vector set ($i < n$), the first i base vectors are always the best possible (in a mean square sense) co-ordinate system to describe the original vectors. The calculation of SVD is quite a common method in linear algebra and is often implemented in specialised software (e.g. Matlab) as a standard function.

The singular value decomposition was applied to the matrix 60x253 constructed from the speed profile vectors' co-ordinates. The new co-ordinate system U contained 60 base vectors and each profile was described by 60 co-ordinates contained by the $S \bullet V^T$ matrix. However, the co-ordinates decreased rapidly from the first towards the last one, which proved that the most characteristic information on the speed profile was contained by the first few co-ordinates. Only two first co-ordinates were chosen as an approximation.

The first 25 speed profiles were manually sorted into three categories. Figure 10 shows a two-dimensional plot where each profile is presented by a point defined by the two co-ordinates. It is clearly seen that the three profile types form quite distinctive clusters in the plot space.

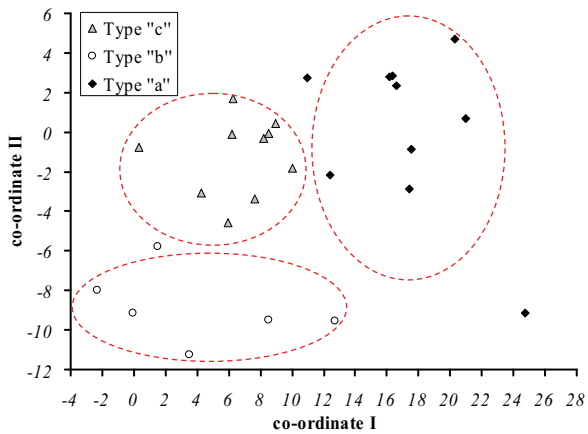
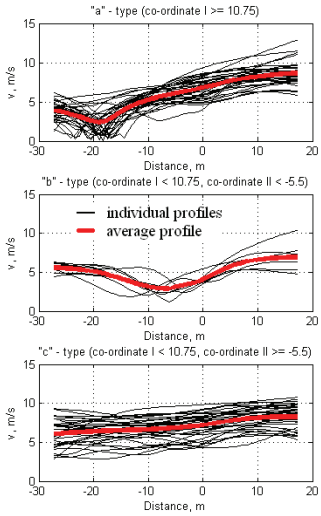


Figure 10. The manual classification of the speed profiles presented by their 2-dimensional approximations.

After the number of dimensions is decreased, the clustering or supervised learning techniques may be used to distinguish the profile shapes. The profiles in Figure ***11*** are split using a simple threshold criteria set for the two co-ordinate values (again, only 75 profiles are plotted).

The classification accuracy of the described technique is dependent on the frequency of the profiles of different shapes. Thus, if a certain type is clearly dominant, it will also "draw" the new co-ordinate system to its side, which may result in the fact that the features of the other, less frequent types, will be lost to a greater extent. Other techniques for dimension reduction and data visualisation that may be tested are, for example, multidimensional scaling (MS, *Borg & Groenen, 2005*) and isomap (*Tenenbaum et al., 2000*).



a)

			SVD situation type "a"	SVD situation type "b"	SVD situation type "c"	Total
Observer situation type		"a"	*36%*	*0%*	*4%*	40%
		"b"	*0%*	*3%*	*3%*	6%
		"c"	*8%*	*0%*	*46%*	54%
Total			44%	3%	53%	100%

Correct classifications - *85%*

b)

Figure 11. The results of profile classification based on 2-dimensional approximations (a) and comparison with the observer's results (b).

4. DISCUSSION AND CONCLUSIONS

There is a great diversity in the road user behaviour forms and classifying it is not a simple task neither for a human, nor for an automated technique. In this paper, we evaluate three pattern recognition techniques according to three criteria: I – quality of the differentiation between the behaviour types; II – use of the variation in the data attributable to different behaviour (profile shapes) and III – ability to treat large datasets.

Criteria II and III are fulfilled by all three methods. The methods utilise the information about profile shapes and can be applied on extensive datasets. When it comes to the criterion I, the quality of the results produced by the three illustrated techniques are not the same.

The *k*-means clustering, which is the least successful, but the simplest one to implement and test, can be recommended when the results have to be produced quickly. The clustering algorithms often assume that there is a set number of clusters in the data and try to "force" the data into them. If the number of clusters is high, but the data is quite homogeneous, the clusters will most probably look alike and the differences might be difficult to interpret. On the other hand, using a higher number of clusters may reveal small peculiar groups not obviously seen among the other profile shapes that are more common. Moreover, more advanced algorithms are capable of making a decision on the optimal number of clusters for the current dataset.

The supervised learning and classification based on the nearest neighbour algorithm shows the best results. The training set reflects human judgements of the expected number of groups and the typical shapes the classification is based on, i.e. it is

exogenous to the data. If a group is missed at the initial stage of classification, it will not be represented in the training set and will be missed in the final classification as well. Another problem is that production of an extensive training set requires much of the manual input.

The dimension reduction techniques are more complex to implement and base on extracting the most typical features from the data and operating the features rather than the original data. This simplifies the analysis and data visualisation to a great extent. For example, drawing up the 2-dimensional approximations of the profiles produced by the SVD algorithm (Figure 10) reveals the pattern of location of different profile types on the graph and allows the setting up of very simple classification criteria that still perform very well. There is a risk, however, that the omitted dimensions may contain important information for the final classification of the profiles and thus affect the quality of the results. As for the quality of the classification, it is still lower compared to supervised learning.

A problem that needs special investigation is the profiles with shapes that do not match any of the typical patterns. All three techniques are quite insensitive to such outliers and simply force them into one of the typical groups. However, examination of the outliers might be important in case they represent some kind of breakdown in normal traffic that might have implication for safety or efficiency. Detailed examination of such situations might give an idea on how they can be eliminated. A possible solution can be to compare individual profiles with the average profile and select significantly different ones.

Finding the right technique for the data is often stated to be more an art than a science, and parameters working well for one dataset will not certainly work for another. The best strategy in this case is to have a toolbox of different techniques where the right one is found by using trials.

The general conclusion is that the pattern recognition techniques perform quite well in classifying the behaviour types compared to classification by a human observer. The advantage of these techniques is the automation of the classification process which allows for analysing large datasets. Another aspect is the reduction of the subjective effects a specific observer might have on the results when doing the classification manually. We have argued in principle that subjective component introduced by an observer might be useful, especially if the differences in behaviour are difficult to express in objective terms. On the other hand, pattern recognition techniques might help reveal the relations between this subjective dimension and objective variables, contribute to standardisation and therefore allow for larger comparability between analyses made by different individuals.

As a direction for future work, pattern recognition techniques can be tested in classifying more complex situations that a human observer manages to classify without being able to explicitly formulate the classification criteria. For example, there is a great interest in finding objective parameters that better reflect the severity of traffic conflicts. This, however, requires a large set of traffic conflicts with detailed data about the road users movements to be available.

5. REFERENCES

Aarts, L., I. van Schagen (2006). Driving speed and the risk of road crashes: a review. Accident Analysis and Prevention 38, pp. 215-224.

André, M., J. Hickman, D. Hassel, R. Jourmard (1995). Driving cycles for emission measurements under European conditions. SAE. Technical Paper 950926.

Ardö, H. (2009). Multi-target tracking using on-line viterbi optimisation and stochastic modelling. Doctoral thesis, Lund University, Faculty of Engineering LTH, Centre for Mathematical Science.

Borg, I., P. Groenen (2005). Modern multidimensional scaling: theory and applications. Springer-Verlag, New York.

Carsten, O. M. J., D. J. Sherborne, J. A. Rothengatter (1998). Intelligent traffic signals for pedestrians: evaluation of trials in three countries. Transportation Research Part C 6, pp. 213-229.

Duda, R. O., P. E. Hart (1973). Pattern classification and scene analysis. A Wiley-interscience publication, New York.

Englund, A., N. P. Gregersen, C. Hydén, P. Lövsund, L. Åberg (1998). Trafiksäkerhet – en kunskapsöversikt (*in Swedish*) Traffic safety - a knowledge review. KFB, Studentlitteratur, Lund.

Ericsson, E. (2000). Urban driving patterns - characterisation, variability and environmental implications. Doctoral thesis, Lund University, Lund Institute of Technology, Department of Technology and Society.

Ericsson, E. (2000). Variability in urban driving patterns. Transportation Research Part D 5, pp. 337-354.

Grayson, G. B. (1984). The Malmö study: a calibration of traffic conflict techniques. SWOV. Report R-84-12.

Hagring, O. (2000). Framkomlighet i korsningar utan trafiksignaler, en litteraturöversikt (*in Swedish*) Traffic efficiency of non-signalised intersections, a literature overview. Lunds Tekniska Högskola, Institutionen för Teknik och Samhälle, Lunds Universitet. Bulletin 190.

Hakkert, A., V. Gitelman, E. Ben-Shabat (2002). An evaluation of crosswalk warning systems: effects on pedestrian and vehicle behaviour. Transportation Research Part F 5, pp. 275-292.

Hydén, C. (1987). The development of a method for traffic safety evaluation: the Swedish traffic conflict technique. Doctoral thesis, Lund University, Department of Traffic Planning and Engineering.

Hydén, C., A. Várhelyi (2000). The effects on safety, time consumption and environment of large scale use of roundabouts in an urban area: a case study. Accident Analysis and Prevention 32, pp. 11-23.

Karlgren, J. (2001). Bilars hastighet längs gator med gupp (*in Swedish*) Vehicles' speed along streets with humps. Doctoral thesis, Chalmers Tekniska Högskola, Göteborg, Tema Stad & Trafik, Sektionen för Arkitektur.

Kloeden, C. N., A. J. McLean, V. M. Moore, G. Ponte (1997). Travelling Speed and the Risk of Crash Involvement. Volume I: findings. University of Adelaide, Road Accident Research Unit. Report No. CR 172.

Larsson, H. (2009). Förarstöd för lägre bränsleförbrukning och minskade emissioner: utvärdering av två system (*in Swedish*) Driver support for better fuel consumption and lower emissions: evaluation of two systems. Licenciate thesis, Lund University, Department for Technology and Society, Faculty of Engineering LTH.

Laureshyn, A., H. Ardö, T. Jonsson, Å. Svensson (2009). Application of automated video analysis for behavioural studies: concept and experience. IET Intelligent Transport Systems 3 (3), pp. 345-357.
Malkhamah, S., M. Tight, F. Montgomery (2005). The development of an automatic method of safety monitoring at Pelican crossings. Accident Analysis and Prevention 37, pp. 938-946.
Martinez, K. L. H., B. E. Porter (2006). Characterizing red light runners following implementation of a photo enforcement program. Accident Analysis and Prevention 38, pp. 862–870.
Messelodi, S., C. M. Modena, M. Zanin (2004). A computer vision system for the detection of vehicles at urbam road intersections. Instituto Trentino di Cultura (ITC). Technical Report T04-02-07.
Michon, J. A. (1985). A critical view of driver behaviour models: what do we know what should we know? *in* L. Evand and R. C. Schwing (Eds.). Human behaviour and traffic safety. Plenum Press, New York**.**
Nilsson, G. (2004). Traffic Safety Dimensions and the Power Model to Describe the Effect of Speed on Safety. Doctoral thesis, Lund University, Department of Technology and Society, Traffic Engineering.
Nygård, M. (1999). A method for analysing traffic safety with help of speed profiles. Master thesis, Tampere University of Technology, Department of Civil Engineering.
Parkhurst, D. (2006). Using digital video analysis to monitor driver behaviour at intersections. Center for Transportation Research and Education, Iowa State University. CTRE Project 05-214.
Pau, M., S. Angius (2001). Do speed bumps really decrease traffic speed? An Italian experience. Accident Analysis & Prevention 33, pp. 585-597.
Peräaho, M., E. Keskinen, M. Hatakka (2003). Driver competence in a hierarchial perspective: implications for driver education. University of Turku, Traffic Research. Report for Swedish Road Administration.
Ripley, B. D. (1996). Pattern recognition and neural networks. Cambridge University Press.
Sekine, T., E. Sekine (2009). Characteristic classification of driving behaviour with mixed traffic situation for driver assist. 21st International Symposium on Dynamics of Vehicles on Roads and Tracks, Stockholm.
Strang, G. (1986). Introduction to applied mathematics. Wellesley-Cambridge Press, Wellesley, Massachusetts.
Tenenbaum, J. B., V. de Silva, J. C. Langford (2000). A Global Geometric Framework for Nonlinear Dimensionality Reduction. Science 290 (5500), pp. 2319-2323.
Várhelyi, A. (1998). Driver's speed behaviour at a zebra crossing: a case study. Accident Analysis and Prevention 30 (6), pp. 731-743.

Paper IV

Published in IET Intelligent Transport Systems
Received on 29th August 2008
Revised on 18th February 2009
doi: 10.1049/iet-its.2008.0077

Special Issue – selected papers from AATT 2008

ISSN 1751-956X

Application of automated video analysis for behavioural studies: concept and experience

A. Laureshyn[1] *H. Ardö*[2] *Å. Svensson*[1] *T. Jonsson*[1]

[1]*Department of Technology and Society, Faculty of Engineering LTH, University of Lund, Sweden*
[2]*Centre for Mathematical Sciences, Faculty of Engineering LTH, University of Lund, Sweden*
E-mail: aliaksei.laureshyn@tft.lth.se

Abstract: Lund University, Sweden, is developing a video analysis system for making long-term behavioural studies, primarily in complex urban environments. Road users are detected using the KLT (Kanade-Lucas-Tomasi) interest point tracker. Trajectories are estimated using foreground–background segmentation, whereas speed is estimated using the shape analysis of interest points. The extracted trajectories are further used for behavioural analysis. The authors present the experience from an ongoing study in Stockholm city, where the task was to find out if allowing two-way bicycle traffic on one-way streets had negative effects on safety. The video analysis system was applied to detect biking in the 'wrong' direction and analyse traffic conflicts between cyclists and other road users. The manual observations done in parallel allowed validating the accuracy of system performance.

1 Introduction

Video recording is commonly used when making behavioural studies in road traffic. It allows collecting data over long time periods and gives the possibility of looking through the material later and in more comfortable conditions. An installed video camera has a less distractive effect on road users compared to a roadside observer. The other advantage is that the video is very illustrative and once an event of interest has been detected it can be looked at again, and additional relevant information can be retrieved. It is also possible to make more exact measurements from video data, for example extract road users' position and speed.

However, as the frequency of searched events decreases and the observational period increases, the problem of event detection becomes crucial. It is resource-consuming to use an observer watching through the entire film, and the results might still be quite inaccurate because attentiveness decreases with time. The extraction of the position and speed data manually is also very time-consuming (e.g. [1] mentions the ratio between the time spent on manually 'clicking' the vehicle position and the length of the video film processed as 10:1).

A research group at Lund University, Faculty of Engineering LTH, is developing a system that can process video data automatically. The system is aimed at studying the behaviour of road users in a complex traffic environment (primarily urban conditions and mixed traffic modes). Now, at the final stage of a four-year project, the system can detect and track vehicles, cyclists and pedestrians and measure their speed and acceleration with relatively high accuracy. Special procedures are developed for describing the interaction between road users, for example detection of encounters (simultaneous presence in a certain area), presence of a collision course, detection of conflict situations, analysis of the speed profile shapes and so on.

This paper describes the methodology implemented in the system on an example of a study where it was practically applied. The study concerns the effects of the introduction of two-way cycling on streets with one-way traffic in Stockholm city and includes observations done in two stages, before and after the introduction. So far only the before-observations have been completed (the second stage is planned for autumn 2009), and therefore it is not possible to draw conclusions about the effects of the

© The Institution of Engineering and Technology 2009

measure. The main focus of this paper lies in the performance of the system, the problems encountered and the perspectives for further improvements.

2 Stockholm study: background and scope

The city of Stockholm has traditionally had a rather small mode share of bicycle trips compared to many other Swedish cities. However, there are initiatives that are trying to promote cycling. One possibility is to allow biking against one-way traffic. This would extend the available network for bicycle trips and lead to shorter travel routes and times. The downside would be that it might also lead to dangerous situations and conflicts between cyclists traveling against one-way traffic and other road users.

The project aims at investigating the total safety effect of allowing cycling against one-way traffic, and not only estimating the risk at a specific street. Therefore the design of the study also includes studying changes in the total bicycle flow and in the route choice, that is from which streets bicycle traffic transfers to the one-way streets. To be able to both establish risk at specific sites and route changes, the sites were in some cases chosen so as to be able to count cyclists at several alternative routes.

Initially, 32 places were selected as potentially interesting for observations. However, finding a good place for camera installations turned out to be a complicated task. The cameras were normally attached to railings on balconies of apartments or offices, but in some places there were no buildings with balconies located near enough. At other sites there were potential camera positions but no electrical power available. Some owners of the buildings did not co-operate, or it was impossible to contact them. Finally, only 22 places were filmed, of which 18 were further analysed. Three of the excluded sites did not have any one-way streets entering or exiting the intersection (only being selected for counting bicyclists), and the fourth was excluded because the camera was too far away from the intersection to allow for proper analysis.

Eight camera units were used for the study (Fig. 1), moved between sites just before or after the weekend and resulting in three to four weekdays of recordings at each site. A camera unit consists of a camera house that contains the camera itself and a mini-server with a high-capacity hard disk (750 Gb). The disk capacity allows continuous filming for approximately 2 weeks with 320×240 resolution and 30 fps frame rate. The unit needs a power supply (12 V) and can be connected to a portable computer via a USB port for adjustment of the camera focus and starting/ending the recordings. A video is stored as 0.5 h files in mjpg format. Further, the video material was processed and the objects moving in the 'wrong' direction were detected. To ensure the quality and also validate the work of the video analysis system, much work is still done manually. This includes (a) calculation of the vehicle, pedestrian and cyclist flows for short periods at each site (b) visual control and sorting of system detections, detecting among them situations that potentially might lead to conflicts.

Figure 1 *Camera unit*

3 Automated analyses of video data

A great deal of progress has been made in constructing systems that can monitor highways where it is only of interest to study motor vehicles [2–6]. Today, there are commercial applications available that can generate trajectories from a highway section where all road users are travelling in the same direction. There are also classical solutions available that can track pedestrians in environments with only pedestrians present, such as parks or walkways [7–9].

Systems that can handle environments where more than one type of road user is present [10, 11] become more advanced as the system now has to determine the type of road user as well. Most systems [3, 6–9, 11, 12] have to be configured for each type of road user. This is typically done by manually specifying a large set of length parameters of some wire frame model or by training the system on a large amount of manually classified training examples. Other methods [4, 10] work with more coarse models where it is enough to specify some approximate size of road users.

When it comes to vulnerable road users, efforts have been made on detecting pedestrians. The current state-of-the-art uses techniques such as bag-of-words [11] (that only uses information from some specific interest points), gradient histogram [13] (that samples the observed image more densely) or randomised forests [14] (that combine the results from several randomly generated decision trees).

Less work has been done on detecting bicycles, and for many of the approaches [14, 15] results are only provided for test images produced from within the scene (such as in

© The Institution of Engineering and Technology 2009
IET Intell. Transp. Syst., 2009, Vol. 3, Iss. 3, pp. 345–357
doi: 10.1049/iet-its.2008.0077

the dataset [16]) and not for the typical surveillance angle where the scene is viewed from above.

Analysis of the digital video records in this study employs several techniques that vary in degree of automation, complexity and computation intensity. Generally, the more advanced the technique, the more sensitive it is to eventual errors, quality of input data and calibration procedures and the more validation it requires.

A 'wrong-way' detector is a relatively simple analysis performed to find situations where something (e.g. pedestrian, vehicle, bicycle or just noise) is moving within the road but in the 'wrong' direction. For each of those events a small video clip is saved to allow the events to be analysed in further detail or manually sorted afterwards. Further the tracks of road users can be extracted and their speeds estimated. Finally, the tracks can be analysed to find specific situations, for example encounters or traffic conflicts.

3.1 'Wrong-way' detector

Wrong-way detection is a fast filter that is capable of removing several uninteresting parts in video sequences. Typically, parameters are chosen to ensure that no interesting events are removed, even if this means that quite a few uninteresting events are included.

The filter was implemented using the KLT (Kanade-Lucas-Tomasi, after the names of the developers) interest point tracker [17]. It finds points in the image that are expected to be easy to identify within the following frames. Typically this consists of points with a lot of structure such as corner points or edge junctions. Then it tracks those points over the entire video sequence. As points are lost, new points are chosen to replace the lost ones. Typically a single road user contains several interest points and large road users contain more interest points than small road users. Some results from this tracker are shown in Fig. 2.

The tracks from the KLT tracker are then filtered based on a set of heuristics. The street is manually marked in the sequence, and all tracks outside it are removed. Also, the 'wrong'- direction is manually specified, and tracks with a direction significantly different from it are removed. This discards road users travelling in the allowed directions and crossing the street. Finally all tracks moving too slowly are removed. This discards some slow-moving pedestrians crossing the street at small angles as well as a lot of noise from interest points detected in the background.

The remaining tracks are counted, and if this count exceeds some threshold the situation is considered interesting and a video cut is extracted and stored for later processing.

3.2 Road user tracking

The tracking of road users can be performed either on the video clips produced by the detector described in the previous section or on the entire video film (usually, a video clip has to be at least 1–2 min long to ensure the quality of the background model; therefore, if the number of detections is high and the clips 'overlap', it is easier to process the entire video material directly).

First a probabilistic background–foreground segmentation similar to the one described in [18] is performed. Background–foreground segmentation is a generalisation of the background subtraction method. Several such methods exist and they are all based on the same principle of estimating the background and then deciding which parts of the image currently show the background and which parts show something else, the foreground.

The method used has a background model that consists of only the temporal mean and variance for each pixel. The ideal way of estimating such a model would be to make a recording of the intersection at some time when there are no road users moving around. In this case the estimation of the background model would simply be the mean and variance estimated at each pixel. Unfortunately, such recordings are never available and for long-term surveillance there might also be changes in the background, which means the background model has to adapt and follow changes in the scene that are significantly slower than the road users. In normal situations it will never take a road user more than a few minutes to pass an intersection even if it has to stop and wait for a read light, which means that changes to the

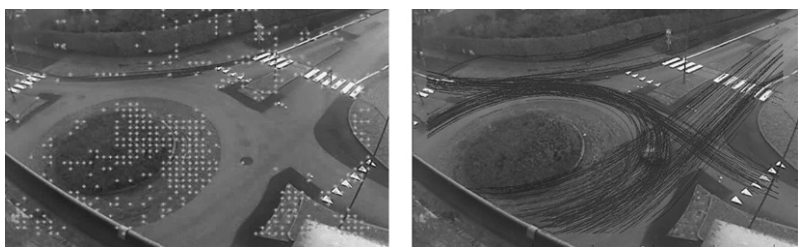


Figure 2 *Results from the KLT interest point tracker [17]. The points on the left picture are the interest points chosen and the lines on the right picture are the tracks generated from a 30 min sequence. Typically each road user generates several tracks*

doi: 10.1049/iet-its.2008.0077
© The Institution of Engineering and Technology 2009

scene that are still in place after 10–20 min could be considered permanent and should be incorporated into the background model.

One solution with such a behaviour is to let the background model be a sliding mean and variance over the last 10 min or so. That requires a 10 min video buffer and is slow because of all the data that have to be processed for each frame. A better solution is to use a learning factor, a number slightly less than one, which indicates how fast the system will forget the history. The background model is then estimated recursively in each frame as the weighted mean of the current background model and the new frame with the learning factor the weight for the background model. This will make the background model a sliding weighted average where the weights decline exponentially with their distance in time from the current frame. This is the classical way of estimating the background model, and it works very well as long as the traffic is not too heavy (because the mean is taken over all frames and not only those showing the background).

In heavy traffic, the estimate can be improved by using the median instead of the mean and by using the 25% and 75% quantiles instead of the variance. As before, slow changes have to be incorporated into the background, which suggests a median filter providing the median over the last 10 min. Such a median can be approximated recursively [19] by incrementing in each frame the background model with some constant amount if it is smaller than the input frame and decrementing it with the same amount if it is larger. In the same way, the 25% quantile can be estimated by letting the decrement used be four times the increment used and the 75% quantile can be estimated by letting the increment be four times the decrement.

Fig. 3 shows a plot of the simulated intensity of a single pixel in grey. In the top-left plot the pixel always shows the

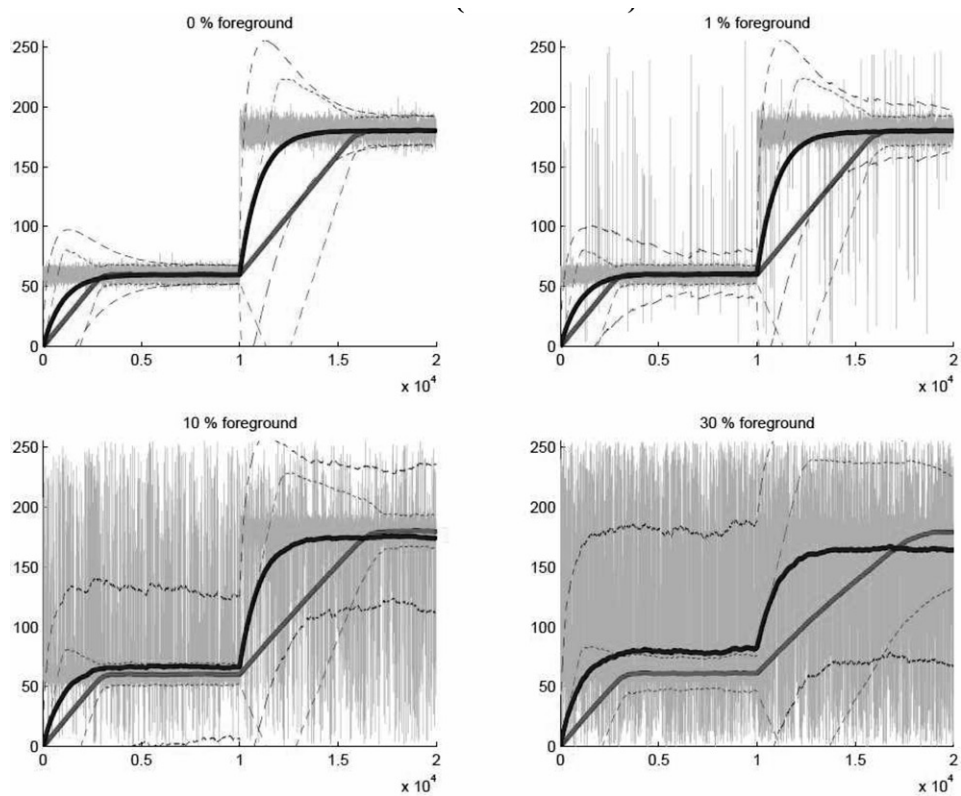


Figure 3 *Comparison of estimations of the mean and variance of normal distributed background process disturbed with different amounts of uniform foreground noise using a quantile-based estimator (black) and a sliding average-based estimator (grey). For both processes the mean value (solid lines) and plus minus two standard deviations (dashed lines) are shown*

© The Institution of Engineering and Technology 2009
IET Intell. Transp. Syst., 2009, Vol. 3, Iss. 3, pp. 345–357
doi: 10.1049/iet-its.2008.0077

background, which is measured with Gaussian noise. In the first half of the sequence the background has intensity 60 and the standard deviation of the noise is 4, and in the second half the background intensity has changed to 180 and the noise level has increased to 6. The blue thick line shows the estimated background model using the first technique based on the mean and variance and the two dashed blue lines show an off-set of two standard deviations from this mean. The red lines show the corresponding values, but are based on the 25%, 50% and 75% quantile estimations instead. Both estimates agree equally well with the ground truth after they have converged. The learning factor and the step size were chosen to make the convergence time of the two estimates approximately equal.

In the top-right image the pixel is assumed to show the foreground 1% of the time. The foreground is modelled as uniformly distributed between 0 and 255. The quantile-based estimator still gives the same result whereas the mean-based estimator over-estimates the variance. In the bottom row the amount of foreground is increased even further and now the mean-based estimator starts over-estimating the background intensity when it is lower than 127 and under-estimating it when it is higher than 127, where as the quantile-based estimator still gives reasonable results. However, the convergence time is increased and in the lower right plot it has not fully converged at the end of the sequence. The estimates archived right before the background intensity was changed and at the end of the sequence are shown in Table 1.

Once the background model is estimated, each 8×8 block of pixels in the model is compared to the corresponding pixels in the current frame using the correlation coefficient. As it is independent of intensity-level scaling and translation, the results become fairly independent on the lighting conditions as long as the lighting is constant over the entire 8×8 patch, that is it fails at the borders of sharp shadows. The output of this step is a probability for each pixel, which is close to 1 if the pixel belongs to the foreground and close to 0 if it belongs to the background. For nearly uniform blocks with very little structure, such as a car roof or road pavement, this probability becomes close to 0.5. This is because in these cases the correlation coefficient becomes unreliable and thus the algorithm becomes more uncertain. Fig. 4 shows some examples of the analysis.

After the foreground probability is calculated for each pixel, the surrounding pixels are used to decide on the final classification. A Markov random field is formulated where in addition to the per pixel foreground probabilities also the pair-wise probabilities of two neighbouring pixels belonging either to the same segment (foreground or background) or to different segments are considered. Typically there is a much higher likelihood for pixels to belong to the same segment than to different segments. By using this information the unknown parts of the image can be filled in as either background or foreground. Also, errors such as small noise segments, small holes or shadow borders are removed.

The problem of solving the Markov random field, for example finding the segmentation that maximises all the probabilities, can be formulated as a graph-cut problem, which can be solved fast and gives a globally optimal solution. This means that among all $2^{640\times480} \simeq 10^{92476}$ possible segmentations for a 640×480 image, the one found is guaranteed to be the one with highest likelihood. In the case of a video it is possible to speed up the calculations even more by utilising the fact that the result for adjacent frames looks very similar; typically the road users have only moved a few pixels each. This is described

Table 1 Comparison of two different techniques to estimate the mean and variance of normal process disturbed with different amounts of uniform noise

Amount foreground (%)	Mean value			Standard deviation		
	True value	Mean estimate	Quantile estimate	True value	Mean estimate	Quantile estimate
0	60	59.93	59.80	4	3.90	3.44
1	60	60.44	59.80	4	9.22	3.44
10	60	66.28	60.04	4	29.88	3.91
30	60	82.52	61.12	4	52.83	7.35
0	180	180.15	179.86	6	5.98	6.61
1	180	179.98	179.86	6	8.34	6.55
10	180	173.77	179.34	6	30.83	7.38
30	180	163.97	178.90	6	48.46	23.69

doi: 10.1049/iet-its.2008.0077
© The Institution of Engineering and Technology 2009

Figure 4 *Background–foreground segmentation example. Left: input frame. Middle: foreground probability per pixel (white pixels have high probability of being a foreground whereas the black pixels have a very low probability). Right: resulting segmentation*

in [20], where they need on average only 33.6 ms to process a 368 × 256 frame and 78.0 ms to process a 720 × 578 frame on a 2.8 GHz P4.

The detected foreground pixels, for example those painted white in Fig. 4 (right), are, within each frame, clustered together into a few connected components, where two foreground pixels are considered connected if they are neighbours. The neighbours of a pixel are defined as the pixels directly above, below, left and right of it, that is each pixel has four neighbours. These connected segments typically correspond to one or more road users. It is only on rare occasions that a single road user is split into several connected components, whereas it is quite common for several road users to belong to the same segments. The latter happens each time one road user is occluded by another, but it is also fairly common for two road users that are very close to become a single connected component, even if there is no occlusion.

These connected components are then used to cluster the interest point tracks from the previous sections into clusters of tracks belonging to the same road user. If there are more interest point tracks connecting two connected components than there are interest point tracks separating them, they are grouped together. This is a very easy operation as there can be no ambiguities, for example either two connected components are in the same group or they are not. There is no optimisation needed. But each group will contain several road users, because whenever there are occlusions road users will be clustered together. To resolve this, the groups have to be split again. As it is fairly uncommon for a single road user to be split into several connected components, events of two sets of tracks belonging to different connected components for more than a few adjacent frames can be used to convert these two sets into two different road users. Allowing a single road user to consist of more than one connected component for a few frames improves the robustness of the system as this does happen for example, when, a car passes behind a lamp post. Fig. 5 shows the tracks from a single car after the clustering is done.

After the clustering of interest point tracks, each road user is represented as a set of interest points in each frame. By calculating the mean over all those points, a point close to the centre of the road user is found, which is used as the resulting position of the generated track. Unfortunately, this position jumps slightly back and forth as interest points are lost and new ones are picked up. This means that it cannot be used to estimate velocities.

Tracks going in the 'wrong' way can be easily detected by using, for example, some entrance and exit gates that a track has to pass. The size of the connected segments representing a road user in the image can give a rough estimation of its real size, which allows one to make estimations on the type of road user (vehicle, pedestrian or cyclist).

3.3 Velocity estimation

To obtain a more precise estimation of the velocity of a road user, a shape analysis of interest points, similar to [21], is performed. The transformation between each set of interest points in terms of rotation translation and scaling is estimated as illustrated in Fig. 6. A mean shape for the total set of interest points is estimated at the same time as the transformations. Then the position in each frame can be expressed as a transformation in terms of rotation translations and scaling from this mean shape into the shape observed in the image. The centre point is then estimated once for the entire track as the mean over all the points in the mean shape. Its position in the different frames is then found by applying the transformation for that frame.

Figure 5 *Interest point tracks belonging to a single road user*

© The Institution of Engineering and Technology 2009
IET Intell. Transp. Syst., 2009, Vol. 3, Iss. 3, pp. 345–357
doi: 10.1049/iet-its.2008.0077

Figure 6 *Transformation between two sets of interesting points*

This gives a smoother estimate of the position as it is no longer assumed that the same set of interest points is available in each frame, so the events of losing interest points and picking up new ones are handled much better. Unfortunately, it is assumed that the shape of the interest points does not change, for example the only transformations it accounts for are rotation translation and scaling. This is not the case when the interest points come from a three-dimensional object viewed with a projective camera, as they can be transformed with a full projective transformation. This means that this measurement is not very good as a position estimate, especially at the beginning and end of the track, where it might be well outside the entire road user. But differentiation of this position gives a very good estimation of the velocity of a road user, whereas measurements from the previous section can be used to estimate the position.

3.4 Systematic errors in speed and position estimation

An image observed in a camera view is a two-dimensional representation of three-dimensional reality; therefore, it is not possible to calculate distances between the objects in reality using measurements in the image only. However, with some prior knowledge about the reality some approximation can be done. For example, if the image is transformed as if taken from straight above the intersection (i.e. rectified), it would resemble an intersection map and the distances can be re-calculated using simple scaling.

The rectified image provides accurate distances between the points in a certain plane in reality, usually the road plane. However, the distances between the object's parts that are elevated would be distorted, and the higher the elevation the higher the distortion. The approximation that the objects are flat and lie in the road plane introduces a systematic error in position estimation as, seen from aside, an object appears to take more place on the road than it actually does. This error depends on the object's height, orientation and distance from the camera, and the angle at which it is seen, that is theoretically the error is not constant as the object passes through the camera view. In practice, the error varies between 1 and 2 m for smaller road users (cars, cyclists and pedestrians), but for larger vehicles, such as buses or lorries, it is much higher.

Despite the error in position the estimation of speed remains quite accurate. If the position error size does not change much during object passage, the speed (i.e. change of the position) is nearly free from error and thus the speed estimations have higher accuracy even without special corrections. Fig. 7 illustrates some results from a test where a car with an installed speed logger was observed and the logged speed was compared with estimations from the video data.

3.5 Detection of conflict situations

Traffic conflicts are situations close to accidents (a breakdown in the interaction between road users) but with a less degree of severity. There are several traffic conflict

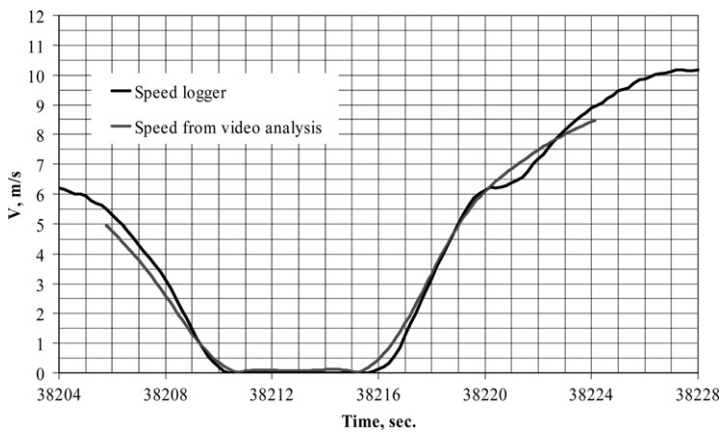


Figure 7 *Speed estimated from video data against speed log in a car*

doi: 10.1049/iet-its.2008.0077
© The Institution of Engineering and Technology 2009

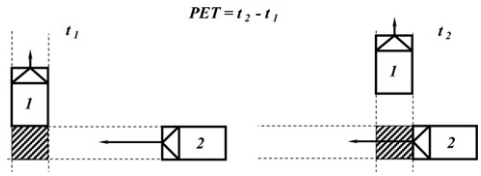


Figure 8 *Conventional definition of PET [1]*

techniques developed, where a conflict is defined using different indicators. The most common indicators are time-to-collision (TTC) and post-encroachment time (PET) or some variation of these parameters [22–24].

TTC is 'the time required for two vehicles to collide if they continue at their present speed and on the same path' [23]. If two road users pass a common spatial zone but at different times, calculation of TTC is not possible. Such encounters can be described by PET, defined as the time difference between the first road user leaving the common spatial zone and the second one arriving at it ([1], Fig. 8). A similar, but continuous, parameter indicating what the PET value will be if road users continue at their present speed and path is called time advantage (TAdv) in [25] and gap time in [24].

The mentioned indicators are quite simple to calculate if two vehicles are on a parallel or perpendicular course (the calculation procedures are, for instance, described in [26, 27]). However, in most general case two vehicles can approach each other from any side and at any angle. Moreover, even for the same approaching angle there are several possible ways for two vehicles to collide (Fig. 9).

It can be seen from the figure that in all the cases it is the corner of one vehicle that hits the side of another. In the most general case one has to separately calculate TTC for a side and a corner of two vehicles in all possible combinations and find the minimal value, which will be the TTC for the vehicles.

The problem in calculating of TAdv is the definition of the common spatial zone, which can be quite large if the crossing angle is not perpendicular (see Fig. 10). In some cases both vehicles can appear in the common zone, but still avoid a collision. To overcome this problem the indicators have to be re-defined in other terms, excluding the uncertain geometrical criteria. Here we used the following definition of TAdv: 'the minimal delay of the first vehicle which, if applied, will result in a collision (assuming that otherwise the vehicles preserve the same speed and path)'. Practically, this also implies that TAdv has to be calculated for all corner-side combinations of the two vehicles and the minimal value will be the TAdv for the vehicles.

Another problem is the definition of the planned path for each road user. A human observer can easily project the planned trajectory, but it is quite difficult to explain how exactly this projection is done. A possible approximation is to assume that the road user actually follows a planned path, that is to calculate indicators along a trajectory that is known. This can be misleading in case the road user avoids a conflict by changing the path, for example taking a larger radius in a turn or changing a lane.

Fig. 11 shows an example of calculated TTC and TAdv profiles for an encounter between two road users. At the beginning (time from T_1 to T_2) they are on a collision course and TTC decreases as they approach each other. However, at the same time vehicle 1 starts braking and vehicle 2 accelerates. Because of this, from moment T_2 they are no longer on a collision course and TAdv starts going up from zero. From moment T_3, vehicle 2 is no longer in the way of vehicle 1 and none of the parameters can be calculated.

PET can be calculated as the time interval between T_3 and the moment when vehicle 1 arrives at the area that has been occupied by vehicle 2 at T_3. However, we consider the TAdv value at T_3 as a more relevant indicator, since PET can be affected by changes in the speed of vehicle 1 after T_3,

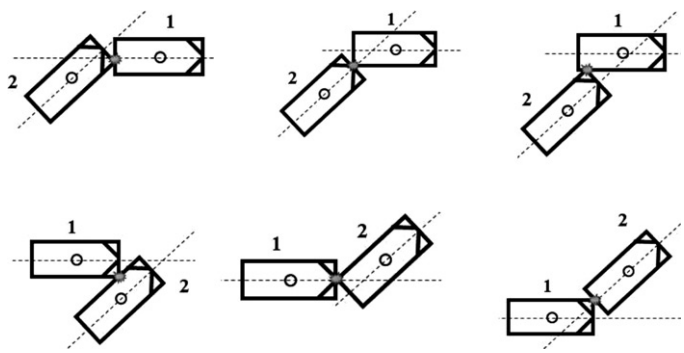


Figure 9 *Possible collision types for the same approach angle (adopted from [27])*

© The Institution of Engineering and Technology 2009
doi: 10.1049/iet-its.2008.0077

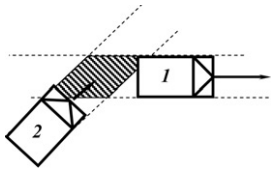


Figure 10 *Problem of calculation of TAdv in case of non-perpendicular approaching angle: the zone becomes larger and the second vehicle can enter it while the first one has not left it and still the collision is avoided*

which might not be really related to the interaction with vehicle 1.

In this way interactions between detected road users can be analysed. By setting certain threshold values for TTC and TAdv, potential conflict situations can be detected.

4 Results

The recording at 18 sites for 3–4 days resulted in 2.5 Tb of data and 900 h of daytime video material. After the first stage of video analysis, that is the detection of objects moving in the 'wrong' direction, this amount was decreased to approximately 27 000 short video clips, with a total length of approximately 115 h. It was decided that this material would be looked through by two observers who sorted all the detections into four categories: bicyclists, pedestrians, cars and other (wrong detections or odd situations). This work took approximately 1 month of full-time work for the observers. The results are presented in Table 2. Since the observational periods were not the same at each site, the numbers are given as an average per day.

To estimate the accuracy of automatic detection, manual cyclist counts were also done at each site for one or two 0.5-h periods (from the video records). The comparison between manual counts and automatic detection is given in Table 3. Initially, manual detections were expected to provide the 'ground truth' with which automatic detection could be compared to. However, at some sites the automatic system found more cyclists than the observers did. Therefore the results in the table are compared to the 'best estimate', which is the maximum between the number of cyclists detected manually and with the help of the video analysis system.

Another task performed by the observers was to detect situations that involve 'wrong-way' cyclists and potentially might lead to conflicts. Totally, only 43 such situations were found, none of which was classified as a serious conflict according to the definition of the Swedish Traffic Conflict Technique [23].

The information available after the video clips had been manually sorted was sufficient for the purpose of the study. Therefore the extraction of road users' tracks was not done on a large scale, but only for a test purpose. Site 33 was chosen for this test as it had a relatively high number of potential conflicts (6) concentrated during four 0.5-h periods (i.e. totally 2 h of video). The tracks were extracted for all road users during this period. Even though there was a possibility to analyse only short sequences detected in the first stage, it

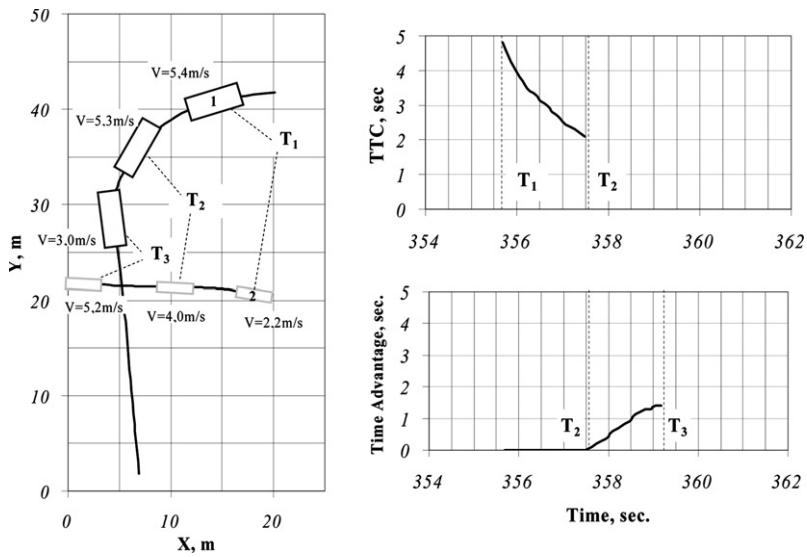


Figure 11 *TTC and TAdv profiles for an interaction – an example*

© The Institution of Engineering and Technology 2009

Table 2 Results of manual sorting of the detections at each site (average per day)

Site	Bicyclists	Pedestrians	Cars	Other	Sum
2	147	894	11	12	1063
4	100	44	19	7	170
5	110	54	9	14	187
6	63	938	26	126	1153
7	8	13	4	1	26
9	42	367	4	159	572
11	35	104	29	63	230
12	13	312	5	5	334
14	31	140	16	48	235
15	35	426	12	26	497
16	55	667	7	17	745
23	208	347	35	54	645
27	52	163	61	61	337
29	13	11	4	14	42
33	55	50	9	6	120
34	28	491	11	18	548
36	30	15	20	10	74
37	12	1	3	4	19
Total	1037	5037	285	645	6997
Percentage	15%	72%	4%	9%	100%

was interesting to compare the performance of these two techniques in the detection of 'wrong-way' cyclists, too. As there are no serious conflicts to be found, the conflict criteria were set quite loose: one of the road users in an encounter had to move in the 'wrong' direction with TTC < 2 s or TAdv < 1 s. Table 4 shows the results of this test.

The studied site appeared in the shade of a large tree during most part of the day. This resulted in many false track detections located on the shade border (as the leaves were moving in the wind, the shades were detected as separate objects). These tracks were, however, very easy to sort out as they were abnormally long in time whereas the travel length did not exceed 1–2 m.

5 Discussion

One of the main benefits of automated video analysis is that it condenses the video material when the events of interest are rare. The amount of raw video data collected in this study is hardly feasible to process employing only human observers. However, the amount of manual work is still significant; therefore, there is an urgent need for further automation of the process (e.g. use of the track detection and analysis technique on a larger scale).

The detection of 'wrong-way' cyclists by the automated system is about 70%, which is quite effective (60 out of 86 in Table 3). However, configuration of the system for a high detection rate results in many false detections. Only 15% of all the detections are cyclists whereas the main part (72%) consists of pedestrians walking on the street (Table 2). This problem can be partly mitigated if a more advanced filter that distinguishes between cyclists and pedestrians is introduced. The filter should include threshold values for both the size and the average speed of a moving object.

Generally, automated detection has a lower detection rate compared to human observers. However, in some cases (sites 2, 5 and 23 in Table 3) the video detector found cyclists that were missed by the observers. All three sites are very lively with pedestrians and cyclists mixed, crossing or moving on the street in all possible directions. Such an environment might be quite distracting for a human observer, whereas the automated detector is not much affected as long as the space between road users is large enough to detect them as separate objects.

© The Institution of Engineering and Technology 2009
IET Intell. Transp. Syst., 2009, Vol. 3, Iss. 3, pp. 345–357
doi: 10.1049/iet-its.2008.0077

Table 3 Manually observed against automatically detected 'wrong-way' cyclists

Site	Time	Best estimate	Manually observed	Automatically detected			
		Bicyclists	Bicyclists (missed)	Bicyclists (missed)	Pedestrians	Cars	Other
2	07.00–07.30	3	2 (−1)	3	7	2	0
4	07.00–07.30	3	3	3	0	0	0
	15.00–15.30	4	4	3 (−1)	0	1	1
5	07.00–07.30	2	2	2	1	0	0
	15.00–15.30	4	2 (−2)	4	3	1	0
6	07.00–07.30	1	1	1	15	3	0
	15.00–15.30	3	3	1 (−2)	29	0	0
7	07.00–07.30	2	2	0 (−2)	0	1	0
	15.00–15.30	1	1	0 (−1)	0	0	0
9	07.00–07.30	2	2	1 (−1)	1	0	0
	15.00–15.30	7	7	4 (−3)	10	0	0
11	07.00–07.30	1	1	1	0	1	0
12	07.00–07.30	0	0	0	2	0	0
14	07.00–07.30	0	0	0	2	1	0
15	07.00–07.30	1	1	1	7	1	1
16	07.00–07.30	3	3	1 (−2)	5	0	1
	15.00–15.30	4	4	0 (−4)	22	0	0
23	07.30–08.00	7	7	6 (−1)	12	0	1
	15.00–15.30	12	7 (-5)	12	18	3	2
27	07.00–07.30	1	1	1	4	1	0
	15.00–15.30	0	0	0	7	0	1
29	07.00–07.30	0	0	0	0	0	0
	15.00–15.30	0	0	0	0	0	0
33	07.00–07.30	8	8	5 (−3)	0	1	0
	15.00–15.30	4	4	4	3	0	0
34	07.00–07.30	1	1	1	9	0	0
	15.00–15.30	5	5	3 (−2)	24	0	1
36	07.00–07.30	1	1	1	0	1	1
	15.00–15.30	3	3	2 (−1)	0	0	1
37	07.00–07.30	3	3	0 (−3)	0	0	0
sum	86	78 (−8)	60 (−26)	181	17	10	

Only four of the six known conflicts were detected by the track-based detector (Table 4). Analysis of the 'misses' showed that in both cases the reason was that the 'wrong-way' cyclists were not detected at all. However, the general detection rates of the simple 'wrong-way' detector and the track-based detector are quite the same (17 cyclists in both cases, but not exactly the same ones), so it might just be a coincidence that the missed cyclists were involved in conflicts. Further tests based on more conflict data will help in obtaining more reliable figures.

© The Institution of Engineering and Technology 2009

Table 4 Detection of 'wrong-way' cyclists and conflicts by two techniques

Period	'Wrong-way' detector (manually classified)			Track-based detection		
	Cyclists	Pedestrians	Conflicts	Cyclists (manually classified)	Pedestrians (manually classified)	Conflicts (automatically detected)
1	8	1	2	9	12	2
2	4	0	1	3	10	0
3	2	3	2	3	5	2
4	3	0	1	2	3	0
total	17	4	6	17	30	4

The systematic bias in position estimation can contribute to the errors in detection. For example, cyclists moving close to the street side might appear as moving on a sidewalk. Extension of the detection zone on a sidewalk, on the other hand, results in many pedestrian detections, which are not relevant (the difference in pedestrian detection can be seen in Table 4, where the detection zone for the track-based detector was extended 0.5 m on a sidewalk compared to the simple 'wrong-way' detector). The quality of the position estimation is also crucial if exact measurements of the conflict parameters (TTC, TAdv) are necessary.

Despite the high number of cyclist passages observed, only a few of them were classified as potential conflicts. It is expected, however, that after the introduction of two-way cycling on one-way streets, cyclist flow will increase and, possibly, the number of conflicts, too. Even if the amount of serious conflicts remains low, analysis of the distribution of TTC and TAdv parameters can give some idea about the changes in safety after the measure introduction.

The installation of cameras turned out to be a complicated task. It is quite difficult to find a place that is close, provides a good view and camera angle, and has power supply. The problem of power supply can be partially solved if mobile sources (batteries) are used. The total power consumption of a camera unit is relatively small and a set of batteries can provide power for approximately one day of observations. However, this introduces new problems of changing the batteries during longer observation periods and securing them from sabotage if left unattended in the open.

A possible way of compensating for a poor camera view and occlusions is to use several cameras looking at a site from different angles. This would allow covering larger areas but would also increase the amount of calculations, and thus the time for analysis, significantly.

At the moment the program code is written in Matlab and partly Visual Basic environments and is not optimised for fastest performance. Therefore there is still the potential to speed up calculations by using C++ language and employing several parallel processors.

Calculation of the conflict indicators relies highly on the accuracy of the position and speed data. The quality of data can be improved not only by using a higher resolution of the video but also by employing data from several cameras, which allow restoring three-dimensional information and thus avoiding the described systematic error. If this problem is solved, the system opens up completely new possibilities for the validation and enhancement of the conflict techniques, which at the present are limited and simplified to match the capacity of human observers used for conflict detection.

6 Conclusions

The developed automated video analysis system has great potential for use in behavioural studies, especially when the studied events are rare. However, at the moment, the amount of false detections is still very high and more advanced filtering algorithms are needed.

It is rarely possible to find a good location to install a camera. Therefore more efforts should be put on developing techniques to compensate for a poor view by using data from several cameras.

There is still a need to improve the accuracy of position and speed estimation, which are important parameters for calculation of safety-related indicators. Since video analysis provides continuous description of road users' trajectories and speed profiles, it provides data for more comprehensive analysis of the behaviour and interactions and can be used, among others, for validation and enhancement of the conflict techniques.

7 References

[1] ARCHER J.: 'Indicators for traffic safety assessment and prediction and their application in micro-simulation

© The Institution of Engineering and Technology 2009
doi: 10.1049/iet-its.2008.0077

modelling'. Doctoral thesis, Royal Institute of Technology, Stockholm, Department of Infrastructure, 2005

[2] KOLLER D., DANIILIDIS K., THÓRHALLSON T., NAGEL H.-H.: 'Model-based object tracking in traffic scenes' in SANDINI G. (ED.): 'Computer Vision–ECCV'92' (Springer, Berlin, Heidelberg, 1992)

[3] TAN T.N., SALLIVAN G.D., BAKER K.D.: 'Model-based localisation and recognition of road vehicles', *Int. J. Comput. Vis.*, 1998, **27**, (1), p. 525

[4] SONG X., NEVATIA R.: 'Detection and tracking of moving vehicles in crowded scenes'. IEEE Workshop Motion and Video Computing, 2007

[5] ZHANG L., LI Y., NEVATIA R.: 'Global data association for multi-object tracking using network flows'. IEEE Conf. Computer Vision and Pattern Recognition, June 2008

[6] SCHOENEMANN T., CREMERS D.: 'Globally optimal shape-based tracking in real-time'. IEEE Conf. Computer Vision and Pattern Recognition, Anchorage, Alaska, June 2008

[7] ISARD M., MACCORMICK J.: 'BraMBLe: a Bayesian multiple-blob tracker'. IEEE Int. Conf. Computer Vision, 2001

[8] ZHAO T., NEVATIA R.: 'Tracking multiple humans in complex situations', *Pattern Anal. Mach. Intell. IEEE Trans.*, 2004, **26**, (9), pp. 1208–1221

[9] BERCLAZ J., FLEURET F., FUA P.: 'Robust people tracking with global trajectory optimization'. Conf. Computer Vision and Pattern Recognition, 2006

[10] ARDÖ H.: 'Multi-target tracking using on-line Viterbi optimisation and stochastic modelling'. Doctoral thesis, Lund University, Faculty of Engineering LTH, Centre for Mathematical Science, 2009

[11] LEIBE B., SCHINDLER K., CORNELIS N., GOOL L.V.: 'Coupled object detection and tracking from static cameras and moving vehicles', *Pattern Anal. Mach. Intell. IEEE Trans.*, 2008, **30**, (10), pp. 1683–1698

[12] KOLLER D., DANILIDIS K., NAGEL H.H.: 'Model-based object tracking in monocular image sequences of road traffic scenes', *Int. J. Comput. Vis.*, 1993, **10**, (3), pp. 257–281

[13] DALAL N., TRIGGS B.: 'Histograms of oriented gradients for human detection'. IEEE Conf. Computer Vision and Pattern Recognition, June 2005

[14] MOOSMANN F., NOWAK E., JURIE F.: 'Randomized clustering forests for image classification', *Pattern Anal. Mach. Intell. IEEE Trans.*, 2008, **30**, (9), pp. 1632–1646

[15] AGARWAL A., TRIGGS B.: 'Hyperfeatures: multilevel local coding for visual recognition'. The Ninth European Conf. Computer Vision (ECCV), 2006

[16] EVERINGHAM M., ZISSERMAN A., WILLIAMS C.K.I., VAN GOOL L.: 'The PASCAL visual object classes challenge', 2006, Available at http://www.pascal-network.org/challenges/VOC/voc2006/results.pdf, accessed February 2009

[17] SHI J., TOMASI C.: 'Good features to track'. *Comput. Vis. Pattern Recognit.* Proc. IEEE Computer Society Conf. CVPR, 1994, pp. 593–600

[18] ARDÖ H., ÅSTRÖM K.: 'Bayesian formulation of image patch matching using cross-correlation'. *Comput. Vis. Pattern Recognit.* Proc. IEEE Computer Society Conf. CVPR 08, 2008

[19] MÖLLER E., GRIESZBACH G., SCHACK B., WITTE H., MAURIZIO P.: 'Statistical properties and control algorithms of recursive quantile estimators', *Biometr. J.*, 2000, **42**, pp. 729–746

[20] KOHLI P., TORR P.H.S.: 'Effciently solving dynamic Markov random fields using graph cuts'. *Comput. Vis.* Tenth IEEE Int. Conf. ICCV, 2005, pp. 922–929

[21] ÅSTRÖM K., KARLSSON J., ENQUIST O., ERICSSON A., KAHL F.: 'Automatic feature point correspondences and shape analysis with missing data and outliers' in ERSBÖLL B.K., STENSTRUP K.S. (EDS.): 'Lecture notes in computer science' (Springer, 2007)

[22] KRAAY J.H., VAN DER HORST A.R.A.: 'The Trautenfels study: a diagnosis of road safety using Dutch conflict observation technique DOCTOR'. Institute for Road Safety Research SWOV, The Netherlands, R-85-53, 1985

[23] HYDÉN C.: 'The development of a method for traffic safety evaluation: the Swedish traffic conflict technique' (Department of Traffic Planning and Engineering, Lund University, Bulletin 70, 1987)

[24] GETTMAN D., HEAD L.: 'Surrogate safety measures from traffic simulation models'. U.S. Department of Transportation, Federal Highway Administration, Report No. FHWA-RD-03-050, 2003

[25] HANSSON A.: 'Studies in driver behaviour, with applications in traffic design and planning'. Two examples' Lund Institute of Technology, Department of Traffic Planning and Engineering, University of Lund, Bulletin 11, 1975

[26] VOGEL K.: 'A comparison of headway and time to collision as safety indicators', *Accident Anal. Prev.*, 2003, **38**, pp. 427–433

[27] VAN DER HORST R.: 'A time-based analysis of road user behaviour in normal and critical encounters', Doctoral thesis, Delft University of Technology, 1990

© The Institution of Engineering and Technology 2009

Paper V

Submitted to "Accident Analysis and Prevention" on 26th September 2009

CYCLISTS IN ROUNDABOUTS – DIFFERENT DESIGN SOLUTIONS

Lisa Sakshaug, Aliaksei Laureshyn[1], Åse Svensson, Christer Hydén

Traffic and Roads, Department of Technology and Society,
Faculty of Engineering LTH, Lund University

Box 118, 22100 Lund, Sweden Tel. +46 46 222 91 25 Fax +46 46 12 32 72

Abstract. Whether the safest roundabout design for cyclists is to separate cycle crossings or integrate cyclists with motorists is an extensively discussed issue. Studies using accident statistics indicate that a separated cycle crossing is the safest for high motor vehicle volumes. However, the results have not been satisfyingly explained. This article combines quantitative and qualitative methods in traffic conflict, interaction and behavioural studies to find out how interactions and conflicts differ between the two roundabout designs. Automated video analysis is used as one of the methods and its performance is evaluated. The integrated roundabout turns out to be more complex with a higher number of serious conflicts and interaction types. The most dangerous situations in the integrated roundabout seem to come about when a motorist enters the roundabout while a cyclist is circulating and when they are both circulating in parallel and the motorist exits. The yielding rules are more ambiguous in the separated roundabout, contributing to a lower yielding rate to cyclists and a lower trust in the other road user's willingness to yield. Situations in the separated roundabout with the lowest yielding rate to cyclists occur when the motorist exits the roundabout at the same time as cyclists are riding in the circulating direction and hence coming from the right. However, most of the accidents in separated roundabouts occur while cyclists are riding against the circulating direction, both when motorists enter and exit the roundabouts.

Keywords: Cyclist safety, roundabout design, behaviour observation, video analysis

1. INTRODUCTION

Although converting an intersection into a roundabout has been shown to result in fewer injury accidents for both motor vehicle drivers and pedestrians (*Elvik & Vaa, 2004, Hydén & Várhelyi, 2000, Schoon & van Minnen, 1994*), the effect on cyclists' safety is unclear or even negative. A Belgian study finds that roundabouts increase cyclist injury accidents by 27 % and fatal accidents by 41-46 % (*Daniels et al., 2008*) whereas a study from the Netherlands shows a 30 % reduction in causalities (*Schoon & van Minnen, 1994*). Danish analyses of 5 years of accidents at roundabouts indicate that cyclists constitute the largest number of accident victims in urban areas

[1] Corresponding author. E-mail: aliaksei.laureshyn@tft.lth.se, tel. +46 46 222 91 31

(*Jørgensen & Jørgensen, 2002*). A Swedish study of large-scale implementation of roundabouts suggests that cycle safety increases and that cyclists are also better respected by motor vehicle drivers (*Hydén & Várhelyi, 2000*).

There are also different views on the safety benefits of special cycle facilities at roundabouts, such as painted cycle lanes, separated cycle crossings and no cycle facilities. Cycle lanes have been shown to be the least safe measure (*Schoon & van Minnen, 1994*). Separated cycle crossings seem to be the safest in situations of high motor vehicle volumes (*Brüde & Larsson, 1999, Schoon & van Minnen, 1994*). For motor vehicle volumes below 8000 incoming vehicles per day, the difference is less clear (*Schoon & van Minnen, 1994*). However, studies of cycle measures in general show that on-road cycling is just as safe or safer than cycling on separated cycle crossings (*Elvik & Vaa, 2004, Aultman-Hall & Hall, 1998*).

An important consideration for cyclist safety is the evidence that motor vehicle drivers primarily look for other motor vehicles and therefore sometimes fail to see cyclists (*Herslund & Jørgensen, 2003, Räsänen & Summala, 2000*, *1998, Summala et al., 1996*). While the aim of integrating cyclists with motor vehicles is to make them more visible, the above-mentioned research shows that it may not be a safer solution in roundabouts. More knowledge is needed to find out what really happens in bicycle-motor vehicle interactions at different roundabout designs. With this in mind, we have studied the behaviour of cyclists and motor vehicle drivers at a roundabout with a cycle path (separated) and one without a cycle facility (integrated).

An important factor considered in our study is the yielding behaviour in cyclist-motor vehicle interactions, which is very complex and, as studies have shown, depends only to some extent on the actual yielding rules. *Räsänen & Summala, 2000*, suggest that motor vehicle drivers' yielding behaviour towards cyclists can be used as one indication of whether the driver has seen the cyclist or not. The cyclist's yielding behaviour is also of interest for the safety situation. Many cyclists lack knowledge of the traffic rules (*Möller et al., 2000*), but knowing when motorists should yield can actually cause accidents as many cyclists rely on them to yield (*Summala et al., 1996*).

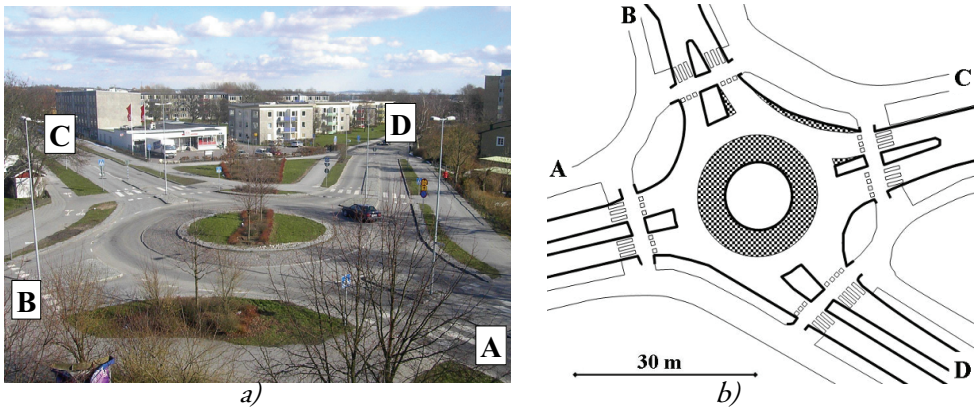
As mentioned above, many quantitative accident studies have attempted to find out which roundabout design is the safest for cyclists but the results have not been satisfyingly explained. The aim of this study is to acquire a better understanding of how road-user behaviour and interactions between cyclists and motorists relate to the safety of cyclists in a separated and an integrated roundabout. This is done by combining quantitative and qualitative studies of traffic conflict types, bicycle-motor vehicle interactions and road users' behaviour. Field studies and video recordings of the two roundabouts are used to analyse traffic conflicts, interactions, yielding behaviour, etc.

2. STUDY SITES

Two roundabouts with similar traffic flows and motor vehicle speeds but different bicycle solutions, one separated and one integrated, were selected for this study. Both roundabouts are situated in Lund, a Swedish town with 100 000 inhabitants and extensive cycling.

2.1. Roundabout designs

The separated roundabout is shown in Figure 1. The cycle paths, together with the pedestrian paths, run parallel to and outside the carriageway. Different surface types are used to distinguish between the pedestrian and cycle paths. Contact between cyclists and motor-vehicle traffic occurs only when a cyclist has to cross the carriageway at a roundabout approach or exit, interacting then with drivers entering or leaving the roundabout. Cycling in both directions on the cycle paths is permitted, which means that drivers have to pay attention to cyclists coming from the left and from the right at the same time. The middle island of the roundabout consists of two parts. The outer ring (≈ 20m in diameter) is paved with cut stones and elevated 5-7cm above the asphalt level. It is quite uncomfortable for cars to drive on it because of the elevation difference, but larger vehicles like trucks and buses can negotiate it without any major inconvenience. The central area (≈ 11m) is further elevated, protected by a curb and not meant for traffic. Sight conditions are very good in all directions.

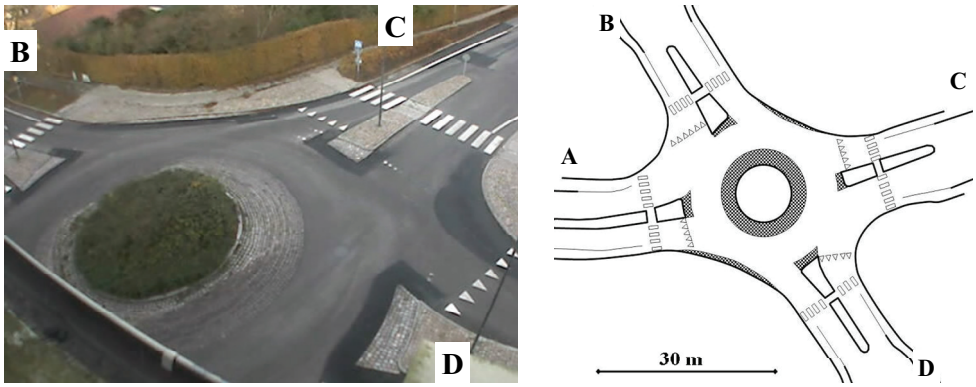


a) b)

Figure 1. Separated roundabout: a) overview photo; b) drawing.

The integrated roundabout is shown in Figure 2. The cycle paths are separated from motor-vehicle traffic along the approach to the roundabout, but cyclists are led onto the carriageway and merged with motor vehicles approximately 30 m before the roundabout. The intention of the design is for cyclists and motor vehicles to form one mixed flow and enter the roundabout and circulate in it as if it was just one lane. However, the widths of the approaches and the ring itself allow cyclists to move in parallel with the vehicles, i.e., two informal lanes are formed. After the roundabout, cyclists are led away from the carriageway again. Cycling is allowed in one direction only on cycle paths along all the approaches, i.e., the cycle path on the right is for those coming towards the roundabout and on the left for those leaving it. The middle island is similar in design to the one in separated roundabout. The outer ring (≈ 14 m in diameter) is paved with cut stones, but with no difference in elevation, which makes it quite comfortable for any type of vehicle to enter. The central area (≈ 9m) is elevated, surrounded by a curb and covered with tight shrubbery. The roundabout is located on a slope (approach C is higher than approach A), which affects the vehicle

speeds, especially cycle speeds. Those coming from approach C have higher speeds, while those coming from approach A have lower speeds compared to approaches B and D. There is a building very close to the roundabout at the corner of approaches A and D, which creates some sight obstruction. The sight conditions at the other corners are quite good.



a) b)

Figure 2. Integrated roundabout: a) overview photo (taken from the balcony of a nearby house, approach A is not seen at all); b) drawing.

2.2. Yielding rules in the separated roundabout

The Swedish yielding rules for cycle crossings at separated roundabouts are unclear and few road users are confident of what applies. In fact, both road users should yield. The cyclist at a cycle crossing should always consider motor vehicles and is only allowed to cross if it can be done safely. The motorist has a stronger yielding obligation when exiting the roundabout than when entering. The entering motorist "should adjust speed in order to not endanger crossing cyclists". The exiting motorist who crosses a bicycle crossing should drive slowly and let crossing cyclists pass (*TRF, 2004*).

2.3. Yielding rules in the integrated roundabout

A cycle is also a vehicle, and therefore all rules concerning vehicles are also valid for cycles. An entering vehicle should yield to circulating vehicles. When the vehicles move in parallel, the roundabout may be seen as having two lanes and thus the vehicle changing lanes or crossing a lane has to yield. When moving in parallel, an exiting motorist should therefore yield to a circulating cyclist (*TRF, 2004*).

3. METHODS

Three different methods are combined in this study: i) field studies; ii) video recording and automated video analysis and iii) accident analysis.

3.1. Field studies

The field studies based on the Swedish Traffic Conflicts Technique (TCT) are used not only to determine the accident risk for cyclists at these roundabouts, but also to study the types of conflict situations and the behaviour preceding conflicts. The TCT uses field observations to study safety based on the relation between serious conflicts and actual accidents. A conflict is defined as "…an observable situation in which two or more road users approach each other in space and time to such an extent that a collision is imminent if their movements remain unchanged." The severity of a conflict is decided based on two variables: Time-to-Accident (TA) and Conflicting Speed (CS). TA is the time from the evasive action until the collision that would occur if the road users' speeds and directions remained unchanged. CS is the speed of the vehicle before the evasive action starts (*Hydén, 1987*).

Two specially trained field observers monitored each roundabout for three days, six hours a day (07:30-9:30, 10:00-12:00, 14:30-16:30) during spring 2008. The conflict observation forms used were modified to suit the specific intersections and included more behavioural information than is usual in traffic conflict studies. One of the observers focused mainly on the cyclists and the other mainly on the motor vehicles. In addition to estimating the speed and the distance to a collision point, they made note of who should yield, which road user passed first, the cyclist's behaviour (*stop, adjust speed, no speed change, get off the bicycle*), the motorist's behaviour (s*top, adjust speed, no speed change*) and, in the integrated roundabout, the behaviour when catching up with another vehicle (*proceeding parallel with the other or staying behind*).

The traffic conflict technique was used in combination with traffic counts from video recordings in order to predict accident risk – accidents per cyclist and accidents per motor vehicle. The data collected during the three days of conflict studies, in combination with the accident analyses from roundabouts in several Swedish cities, also provided a means of characterizing all the typical collision situations for a specific site (*Hydén, 1987*). These typical situations were then used as a basis for studying video recorded behaviour at the two roundabouts in interactions similar to serious conflicts. Interactions are defined as situations between two road users "where at least one of the road users changes speed or direction because of the other". For the integrated roundabout this included situations in which a cyclist and car ended up parallel with each other on the carriageway, where there was also an interaction in most cases, albeit more subtle. Together the TCT, accident analyses and the analysis of the interactions provided information for a thorough description of the cyclists' safety situation at the two roundabouts, enabling interpretation of why and how risks were generated in the two types.

The speeds were measured with a radar gun at one of the approaches (approach C at both roundabouts, Figure 1-Figure 2) for free vehicles entering and exiting the roundabouts and for free cyclists in the middle of the crossing at the separated roundabout and when entering and exiting the integrated roundabout. "Free" is defined as not affected by other road users, which in practice means that there are no other road users closer than 5 seconds ahead. 100 measurements were taken for each

vehicle flow and 50 for each bicycle flow (bicycles going both ways on the crossing at the separated roundabout were considered as the same flow).

3.2. Video recordings and video analysis

Video recordings were performed for five days at each roundabout. The separated roundabout was filmed in the middle of April 2008 and the integrated roundabout in the middle of November 2006. The cameras were mounted on nearby buildings. At the integrated roundabout it was not possible to get a view over the entire area and one of the approaches is not seen at all, while two others are seen only partly. To be able to compare the sites, the corresponding limitations are presumed for the separated roundabout, too (Figure 3).

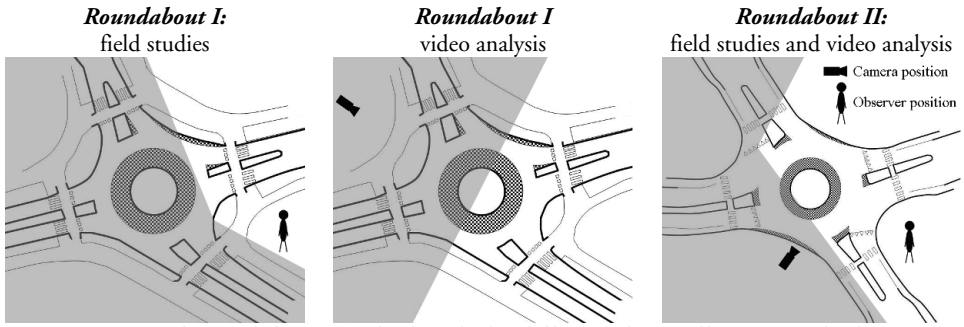


Figure 3. Parts of roundabouts studied with the different data collecting methods.

Three observers manually registered the interactions, behaviour and route choices of all passing cyclists for a period of 24 hours (a weekday). Motor vehicles interacting with cyclists were registered in a similar way. In order to obtain estimates of the traffic flows, all motor vehicles passing during the first five minutes of every hour were also counted. As in the field studies, the road user who passed first in an interaction was noted as "yielded to". It is thus possible that both road users adjusted their speeds or stopped; however, it was always one who acted to finally let the other pass.

The recorded video data (except for the hours of darkness, i.e., 9 hours per day) was processed by the automated video analysis system developed at Lund University (technical details about the system can be found in *Ardö, 2009, Laureshyn et al., 2009*) that detected the cyclists passing at the studied locations. The detection was performed in several steps. First, projected ground-plane tracks of all the moving objects detected in the video were produced. This included the trajectories of the road users but also much noise, such as swinging tree branches, birds and the like. In the second step, only the tracks going in pre-defined directions were selected. To be selected, a track had to pass through a "gate", i.e. cross two lines, "an entrance" and "an exit", in the right order (Figure 4). The definition of the gates was complicated by the fact that the estimated trajectories of road users were systematically shifted from the true position (the problem is discussed in *Laureshyn & Ardö, 2006*); therefore the positions of the gates had to be shifted as well. Another problem was that a trajectory might have been split into several tracks in case the road user was occluded by some other object and lost by the tracking algorithm. As a result, such trajectories might be

missed by the gate-detector. The system utilised some techniques which partly compensated for occlusions. Usually, occlusion by small items of road furniture (e.g. a lamp post) is not a serious problem, at least not for larger objects like cars, lorries and buses. However, if the sizes of the road users in pixels are also small (which is the case with pedestrians and cyclists, moving far from the camera), they nearly "disappear" behind a lamp post and are often lost by the tracking algorithms. At both sites the location of the camera was not optimal with respect to distance from the filmed area and there were lamp posts located in front of the cyclists. At separated roundabout nearly all the cyclist tracks were split. After some experimenting, however, the gate definitions that provided quite satisfactory detection rates were found.

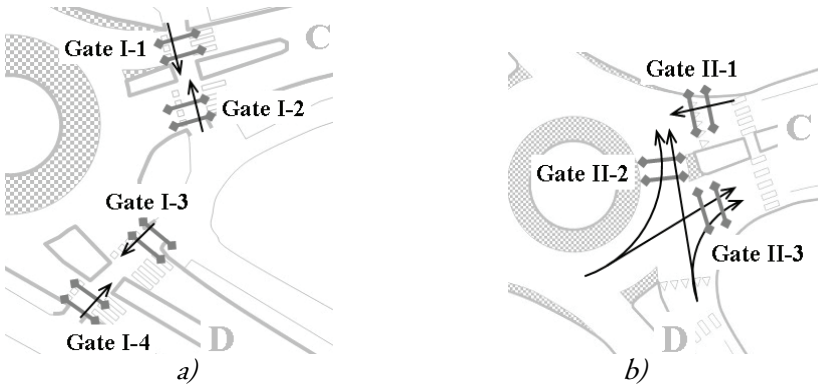


Figure 4. The schematic location of the gates used to detect cyclists at separated (a) and integrated roundabouts (b).

Finally, the cyclists were filtered from other road users moving in the same direction, based on the average size in pixels of the detected object during the passage. This criterion works well for distinguishing between cyclists and vehicles, but is not good enough for distinguishing between cyclists and pedestrians; both are quite similar in size, especially at a distance. For this reason the filter did not perform well at separated roundabout (where cyclists shared the crossings with pedestrians) and was skipped. At integrated roundabout, the camera was located too close to the site, which resulted in some road users appearing at a very short distance from the camera and occupying a large space in the image (in pixels). For these situations the detection algorithms did not perform well, detecting one road user as a group of several smaller ones. This affected the detection results considerably at Gates II-2 and II-3 where a very high percentage of false positive detections (parts of vehicles detected as cyclists) was found.

For each detected cyclist a link, containing information about his/her time position in the original video file, was saved, which allowed quick browsing of the detections to see what was detected. The detections were examined by the observers, who classified the type of the detected object (cyclist, pedestrian/vehicle or other) and also analysed the situations that appeared to be conflicts with cyclists involved.

3.3. Accident statistics

The analyses of the accidents at the two sites involving cyclists are registered in the Swedish Traffic Accident Data Acquisition (STRADA) and contain both hospital and police records. However, since the roundabouts were built/rebuilt recently there is not enough accident data to warrant any firm conclusions. The separated roundabout was built in 2003. From 2003 to May 2008 only one accident with slight injuries involving a cyclist occurred. The integrated roundabout was equipped with a painted bicycle lane in 1999, and in 2006 the bicycle lane was removed and the roundabout was narrowed slightly with cobble stones as extensions of the traffic islands. No hospital- or police-reported bicycle or motor vehicle accidents were reported between the rebuilding and May 2008.

As accident data from the two sites in Lund is too limited to make a better prediction of what generally occurs in the two types of roundabouts, it has been complemented with data from a number of other cities in Sweden, see Table 10.

4. RESULTS

4.1. Traffic flows and speeds

The daily (24-hour) vehicle and bicycle flows at two roundabouts based on the traffic counts from the 24-hour video are shown in Figure 5 and Figure 6.

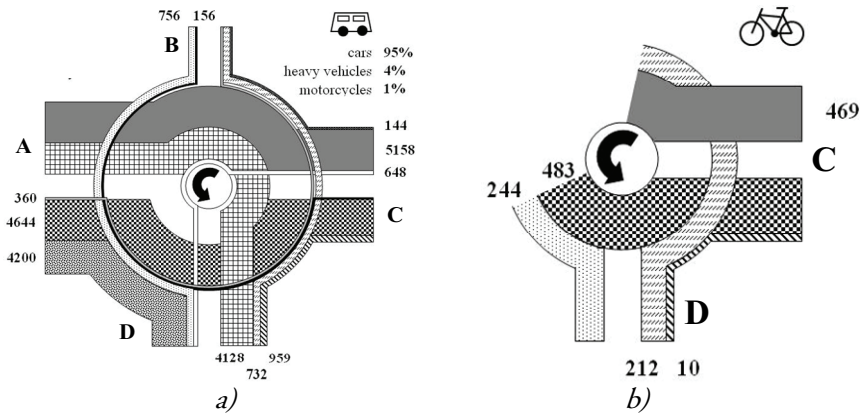
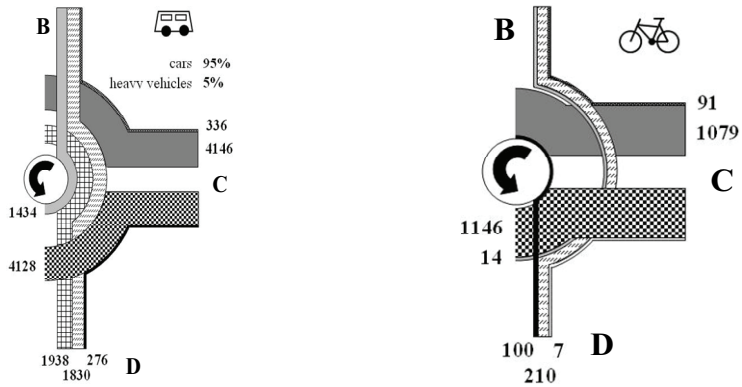


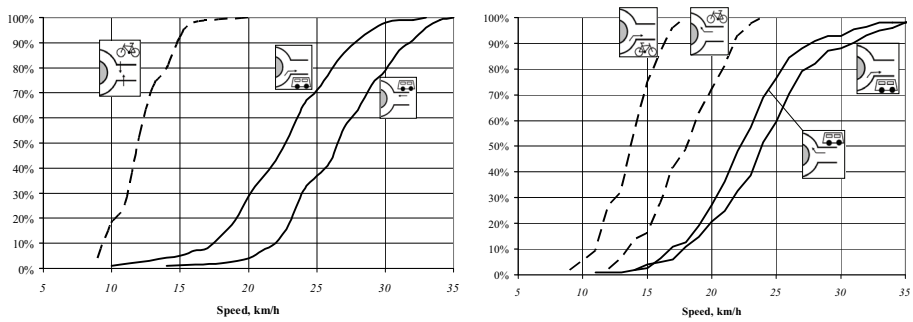
Figure 5. Traffic flows at separated roundabout: a) vehicle daily (24-hour) flow diagram; b) bicycle daily (24-hour) flow diagram (since only approaches C and D were studied, detailed information on approaches A and B was not collected).



a) b)

Figure 6. Traffic flows at integrated roundabout: vehicle (a) and cycle (b) daily (24-hour) flow diagrams (information for approach A is not available and is incomplete for approaches B and D).

The distribution of vehicle and bicycle speeds is shown in Figure 7.



a) b)

Figure 7. Vehicle and cycle speed distributions: a) separated roundabout, approach C; b) integrated roundabout, approach C.

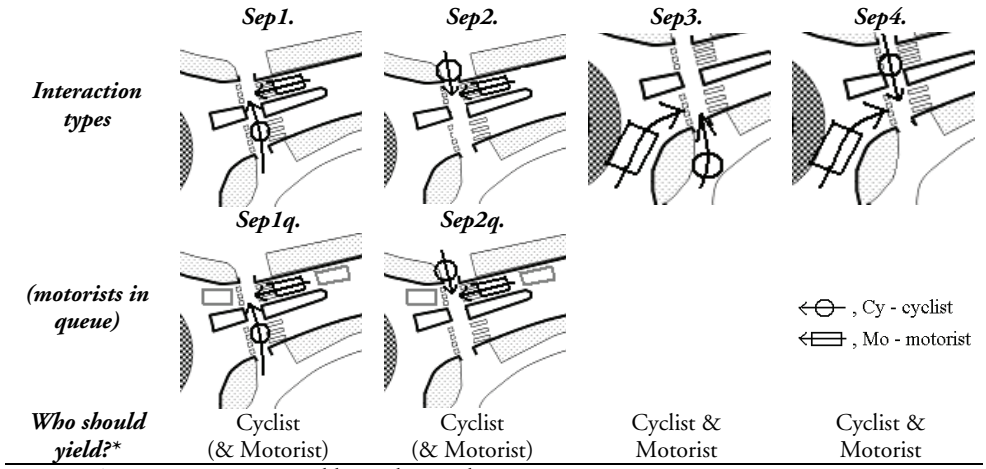
4.2. Interaction types in the separated roundabout

Four main interaction types were observed during the video-recorded 24 hours in the separated roundabout with cycles allowed in both directions. As shown in * For more extensive yielding rules see chapter 2.

Figure 8 they are:

- Sep1 – entering motorist and cyclist in the circulating direction;
- Sep2 – entering motorist and cyclist against the circulating direction;
- Sep3 – exiting motorist and cyclist in the circulating direction;
- Sep4 – exiting motorist and cyclist against the circulating direction.

The field studies also showed that the behaviour differed when motorists were in a queue. These situations are therefore treated separately (Sep1q and Sep2q).



* For more extensive yielding rules see chapter 2.

Figure 8. Interaction types in the separated roundabout.

4.3. Interaction types in the integrated roundabout

Six interaction types were observed from the video-recorded 24 hours in the integrated roundabout. As shown in * For more extensive yielding rules see chapter 2.

Figure 9 they are:

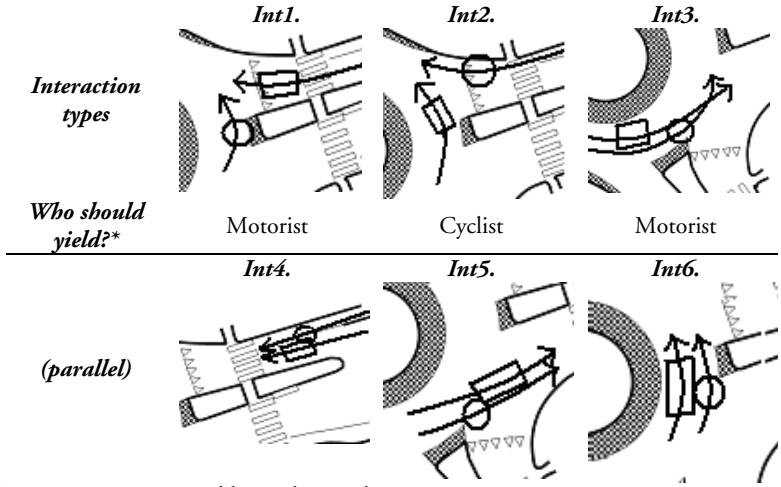
- Int1 – entering motorist and circulating cyclist;
- Int2 – circulating motorist and entering cyclist;
- Int3 – exiting motorist and circulating cyclist;
- Int4 – motorist and cyclist entering in parallel;
- Int5 – motorist and cyclist exiting in parallel;
- Int6 – motorist and cyclist circulating in parallel.

4.4. Serious conflicts in the separated roundabout

During the three days of conflict studies in the separated roundabout there were 2 serious conflicts. This corresponds to 2.3 serious conflicts per 1000 cyclists and 0.2 per 1000 motorists (only entering and exiting motor vehicles included in the traffic count). Traffic conflict studies not only provide a prediction of the expected number of accidents at a certain intersection but also give information on the type of accidents (*Hydén, 1987*). Hence, the qualitative description of the serious conflicts should cover the expected accident types in the roundabouts, even though estimation is poor due to small numbers.

In both serious conflicts the cyclists did not change speeds and seemed to assume that the motorist would yield. In one of the cases (Sep1q) the motorist was in a queue and started driving just before the cyclist, who was approaching at an estimated speed of 16 km/h, was about to pass. The motorist was probably focusing on the motor vehicles on the roundabout and in front of him and therefore observed the cyclist late.

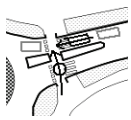
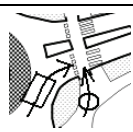
The cyclist swerved and both stopped. In the other case (Sep3) the cyclist continued at unchanged speed and the motorist, approaching from the roundabout, braked hard, probably because he noticed the cyclist late.

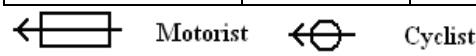


* For more extensive yielding rules see chapter 2.

Figure 9. Interaction types in integrated roundabout.

Table 1. Serious conflicts in the separated roundabout.

Interaction type	*Number of conflicts*	*Evasive action*
Sep1q 	*1*	Cyclist swerves, both stop
Sep3 	*1*	Motorist stops

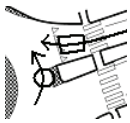
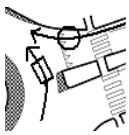
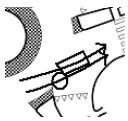


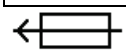
4.5. Serious conflicts in the integrated roundabout

During the three days of conflict studies there were 10 serious conflicts in the integrated roundabout, all but two of them more severe than those in the separated roundabout (according to severity classification by *Svensson, 1998*). This corresponds to 4.6 serious conflicts per 1000 cyclists and 0.2 per 1000 motorists. Four of the serious conflicts were of type Int1; the cyclist was already circulating and the motorist who was entering should therefore have yielded. However, it was still the cyclist who took evasive action in two of the four cases. Three of the conflicts were of type Int2; the motorist was already circulating and the cyclist who was entering should therefore

have yielded. The motorist took evasive action in two of the three cases. In the two serious conflicts of type Int3 the motorist and cyclist circulated in parallel and the motorist, who exited, should therefore have yielded. The motorist took evasive action in both cases but the cyclist also stopped in one of them. In the last serious conflict, of type Int5, a lorry exited in parallel with a cyclist who was almost squeezed. The cyclist took evasive action.

Table 2. Serious conflicts in the integrated roundabout.

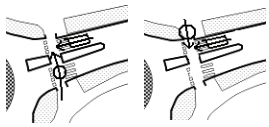
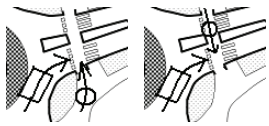
Interaction type	*Number of conflicts*	*Evasive action*
Int1 	*4*	Cyclist stops Cyclist swerves/brakes Motorist brakes Motorist brakes
Int2 	*3*	Motorist brakes Motorist brakes Cyclist brakes
Int3	*2*	Motorist brakes, cyclist swerves/stops Motorist brakes
Int5 	*1*	Cyclist brakes /swerves (motorist is a lorry)

 Motorist Cyclist

4.6. Yielding behaviour in the separated roundabout

The motorists yielded in 68% (n = 354) of the interactions and the cyclist in 32% (n=164) of the interactions at the cycle crossing in the separated roundabout. As shown in Tables 3-5, the motorists yielded to the cyclists to a larger extent when entering the roundabout than when exiting. This difference is smaller when only including non-queuing situations. The motorists entering the roundabout yielded to a larger extent to the cyclists coming from their left (in the direction of the circulation). The exiting motorists also yielded to a larger extent to the cyclists coming from their left (against the direction of the circulation). The queuing motorists yielded to a larger extent than the free vehicles.

Table 3. Yielding of motorists when entering and exiting the separated roundabout.

Interaction type	Total	Motorist yields		Chi2, p-level
		n	%	
Sep1, Sep1q, Sep2, Sep2q (Sep1, Sep2 only) 	277 (216)	207 (149)	75% (69%)	P=0.001 (P=0.078)
Sep3, Sep4. 	241	147	61%	

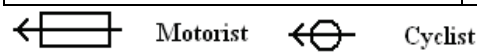
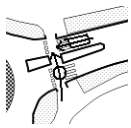
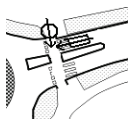
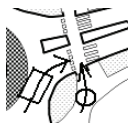
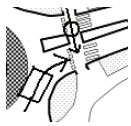


Table 4. Yielding of motorists when cyclists moved in and against the circulating direction (separated roundabout).

Interaction type	Total	Motorist yields		Chi2, p-level
		n	%	
Sep1, Sep1q 	132	115	87%	P=0.000
Sep2, Sep2q 	143	91	64%	
Sep3 	169	93	55%	P=0.003
Sep4 	73	55	75%	

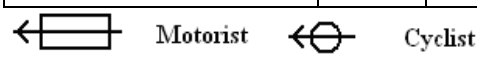
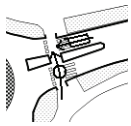
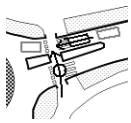
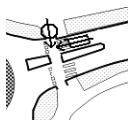
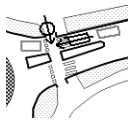


Table 5. Yielding of motorists in non-queuing and queuing situations (separated roundabout).

Interaction type	Total	Motorist yields		Chi2, p-level
		n	%	
Sep1 	102	86	84%	P=0.076
Sep1q 	30	29	97%	
Sep2 	112	62	55%	P=0.000
Sep2q 	31	29	94%	

Motorist Cyclist

4.7. Yielding behaviour in the integrated roundabout

4.7.1. Int1 – cyclist circulating, motorist entering

A vehicle entering a roundabout is obliged to yield to a circulating vehicle, hence the motorist should yield to the cyclist. There were 138 Int1-interactions and the motorist did not yield in 4% (n=6) of those. In four of the six interactions where the motorists did not yield, they did not adjust at all to the circulating cyclists. Even when the motorist yielded, the cyclist also adjusted his speed or direction in 8 % of the interactions.

4.7.2. Int2 – cyclist entering, motorist circulating

The cyclist is entering the roundabout and should therefore yield. There were 171 Int2-interactions. The cyclists did not yield in 14% of those. In 18 of the 24 interactions where the cyclists did not yield, they did not adjust at all to the circulating motorists. In 7 of these 18 interactions the motorists also continued and circulated in parallel, and in the rest of the cases the motorist yielded. Even when the cyclist yielded, the motorist adjusted his speed or direction in 8 % of the interactions.

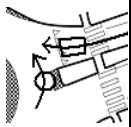
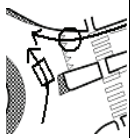
4.7.3. Int3 – cyclist circulating, motorist exiting

The motorist is changing lanes and should therefore yield. There were 23 Int3-interactions. The motorist did not yield, but continued at the same speed, in 6 of

these interactions. In 2 of the 17 interactions where the motorist did yield, the cyclist also adjusted speed or direction.

Table 6. summarises cyclists' and motorists' behaviour in interactions Int1-3.

Table 6. Behaviour in different interaction situations (integrated roundabout).

Interaction type	*Who should yield?*	*Who yields?*			*Behaviour*			
			n	%	Cyclist adjusts	Cyclist unchanged	Motorist adjusts	Motorist unchanged
Int1 	Motorist	Motorist	*132*	*96%*	*11 (8%)*	*121 (92%)*		
		Cyclist	*6*	*4%*			*2*	*4*
Int2 	Cyclist	Motorist	*17*	*10%*	*3*	*14*		
		Cyclist	*147*	*86%*			*12 (8%)*	*135 (92%)*
		Parallel	*7*	*4%*	*3*	*4*		*7*
Int3	Motorist	Motorist	*17*	*74%*	*2*	*15*		
		Cyclist	*6*	*26%*				*6*
		Not known	*4*					

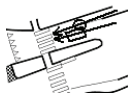
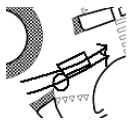
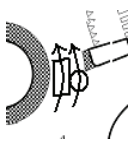
Motorist Cyclist

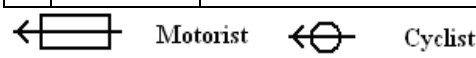
4.7.4. Int4, 5, 6 – motorist and cyclist move in parallel

When a cyclist and a motorist caught up with each other they had a choice to make. They either stayed behind or moved in parallel with each other. Moving in parallel could lead to squeezing or Int3-situations. Table 7 shows the cyclists' and motorists' behaviour when catching up with each other in the integrated roundabout.

More than half of the motorists who caught up with cyclists did not stay behind but moved in parallel. Most of the cyclists stayed behind. The road users who moved in parallel were mainly a problem when it led to Int3 or if the cyclist was squeezed by the motorist. There were too few interactions with heavy vehicles to find more than tendencies. All of the 8 cyclists who caught up with heavy vehicles stayed behind them. However, not all the heavy vehicles stayed behind the cyclists when catching up with them; 6 of 15 moved in parallel with the bicycle.

Table 7. Behaviour in the parallel interaction situations (integrated roundabout).

Interaction type		Who catches up?	Behaviour	All motor vehicles		Heavy vehicles
				n	%	n
Int4		Cyclist	Moves in parallel	18	33%	-
			Stays behind	36	67%	5
			Total	54	100%	
		Motorist	Moves in parallel	40	49%	-
			Stays behind	41	51%	1
			Total	81	100%	
		Parallel all the way through		21		-
		TOTAL		156		6
		MISSING		19		
Int5		Cyclist	Moves in parallel	1		-
			Stays behind	-		-
			Total	1		
		Motorist	Moves in parallel	17	68%	1
			Stays behind	8	32%	-
			Total	25	100%	
		Parallel all the way through		-		-
		TOTAL		26		1
		MISSING		14		
Int6		Cyclist	Moves in parallel	5	17%	-
			Stays behind	25	83%	3
			Total	30	100%	
		Motorist	moves in parallel	168	56%	5
			Stays behind	132	44%	8
			Total	300	100%	
		Parallel all the way through		15		-
		TOTAL		345		16
		MISSING		19		



4.8. Behaviour in the two roundabouts

As shown in Table 8, a larger proportion of the motorists and cyclists interacting with each other continued at an unchanged speed and direction in the integrated roundabout compared to the separated roundabout (Chi2: motorists p=0.003, cyclists p=0.000). In the separated roundabout 22% of the motorists and 9% of the cyclists stopped in interactions compared to 5% and 3% respectively in the integrated roundabout (Chi2: motorist p=0.000, cyclist p=0.000). In addition, a larger proportion of the motor vehicles had already stopped for other motor vehicles, pedestrians or cyclists and were therefore standing still when the interacting cyclist arrived at the separated roundabout compared to the integrated roundabout (Chi2: p=0.000). This was probably partly due to the fact that the cyclist and motor vehicle

in the integrated roundabout could interact without crossing each others' paths, but with just a small change of direction, by moving in parallel instead of yielding.

Table 8 Cyclists behaviour in interactions with motorists in the integrated and the separated roundabout.

Roundabout type	*No speed change*		*Get off the bike*		*Adjust speed*		*Stop*		*Stand still*		*Total*
	n	%	n	%	n	%	n	%	n	%	
Integrated	*545*	*59%*	*1*	*0%*	*339*	*37%*	*25*	*3%*	*7*	*1%*	*917*
Separated	*189*	*36%*	*7*	*1%*	*273*	*52%*	*49*	*10%*	*5*	*1%*	*523*

Table 9 Motorists behaviour in interactions with cyclists in the integrated and the separated roundabout.

Roundabout type	*No speed change*		*Adjust speed*		*Stop*		*Stand still*		*Total*
	n	%	n	%	n	%	n	%	
Integrated	*409*	*45%*	*431*	*47%*	*48*	*5%*	*28*	*3%*	*916*
Separated	*180*	*34%*	*178*	*34%*	*114*	*22%*	*53*	*10%*	*525*

In the separated roundabout 3% of the 1453 cyclists that passed during the 24-hour period cycled on the carriageway. Correspondingly, in the integrated roundabout, 5% of the 2677 cyclists moved onto the footpath and crossed the street at the zebra crossing. In the separated roundabout (with bi-directional cycle path) 38% moved against the circulating direction. In the integrated roundabout 3% moved against the circulating direction. Surprisingly, not all the cyclists who moved against the circulating direction in the integrated roundabout cycled on the zebra crossing; 6% (n=68) cycled onto the carriageway.

4.9. Accident statistics from separated and integrated roundabouts

One can see from the tables that there are some differences among the cities, especially in the separated roundabouts. However, it seems as if the accidents are split in a fairly equal manner between vehicles entering and vehicles exiting. It also seems as if cyclists riding against the circulating direction represent around two thirds of all accidents. Sep2, entering vehicles and cyclists from the right, (against the circulating direction) is the largest individual type (38%).

Table 10 Separated roundabouts. Cycle accidents in nine cities in Sweden for 5 years (2004-2008) presented with regard to type.

City	***Sep 1***	***Sep 2***	***Sep 3***	***Sep 4***	***Unknown /irrelevant***	***Total*** (excluding unknown)
Malmö	*5*	*11*	*7*	*10*		*33*
Gävle		*2*	*2*	*4*		*8*
Linköping	*1*	*5*				*6*
Eskilstuna	*1*	*5*	*1*		*3*	*7*
Västerås	*5*	*3*	*3*	*4*	*1*	*15*
Borlänge	*2*				*1*	*2*
Jönköping				*1*	*2*	*1*
Umeå			*1*	*3*	*1*	*4*
Helsingborg	*2*	*1*	*1*		*3*	*4*
TOTAL, n	*15*	*23*	*19*	*23*	*11*	*47*
TOTAL, %	*13%*	*38%*	*19%*	*30%*	-	*100%*

Table 11 Integrated roundabouts. Cycle accidents in six cities in Sweden for 5 years (2004-2008) presented with regard to type.

City	***Int 1***	***Int 2***	***Int 3***	***Int 4***	***Int 5***	***Unknown /irrelevant***	***Total*** (excluding unknown)
Malmö	*8*	*2*	*1*				*11*
Växjö	*11*		*3*	*1*	*1*		*16*
Linköping	*2*						*2*
Västerås	*1*					*1*	*1*
Borlänge	*3*						*3*
Helsingborg			*1*				*1*
TOTAL, n	*25*	*2*	*5*	*1*	*1*	*1*	*34*
TOTAL, %	*73%*	*6%*	*15%*	*3%*	*3%*		*100%*

In integrated roundabouts by far the most common type is Int 1, motorists entering the roundabout with cyclists circulating, while Int 2, cyclists entering the roundabout with motorists circulating, only occurred in 2 out of 34 accidents. All the other accidents were in different ways due to cyclists and motorists moving in parallel, where the largest problem seems to be linked to Int 3, i.e., motorists leaving the roundabout.

4.10. Detection quality of the automated video analysis system

The use of the automated video analysis system "condensed" the video data from 90 hours down to 35 hours of recordings that were then scrutinized manually by the observers. However, some cyclists were missed by the system, while some of the detections were not cyclists (false positives). Table 12 presents the comparison between the results of the automated detection and the "ground truth" (cyclist counts done by an observer who watched through the same video).

Table 12 Comparison between the automated and manual detections of cyclists (9-hour period).

Cyclist direction		Cyclists, "ground truth"	Automated video analysis				
			Cyclists	Detection rate*	False positives	False positive rate**	Detections, total
Separated roundabout	Gate I-1	149	98	66%	52	35%	150
	Gate I-2	242	104	43%	48	32%	152
	Gate I-3	99	31	31%	20	39%	51
	Gate I-4	177	32	18%	8	20%	40
Integrated roundabout	Gate II-1	541	387	72%	198	34%	585
	Gate II-2	172	82	48%	709	90%	791
	Gate II-3	832	485	58%	893	65%	1378

* Detection rate is calculated as a ratio between the number of correct detection and the "ground truth".

** False positive rate is a ration between the number of false positives and the total number of detections.

The distribution of the automated detections by type over the entire period (5 days) is shown in Table 13.

Table 13. Distribution of the automated detections by type (5 days, 45-hour period)

Cyclist direction		Cyclists	False positives	False positive rate	Detections, total
Separated roundabout	Gate I-1	518	326	39%	844
	Gate I-2	565	328	37%	893
	Gate I-3	123	147	54%	270
	Gate I-4	131	61	32%	192
Integrated roundabout	Gate II-1	1507	556	27%	2063
	Gate II-2	381	2223	85%	2604
	Gate II-3	not available			5369

5. DISCUSSION

This study has shown the difficulties in relating behaviour and interactions to traffic safety. The field studies alone did not give sufficient numbers of conflicts but together with accident data from other cities we have obtained a satisfactory general picture of what seem to be the most important accident types and what distinguishes the integrated roundabout from the separated one. Our results have produced a general

overview of cyclists' interactions with motorists in a separated and in an integrated roundabout. The yielding behaviour differs extensively between and within the roundabouts and we will now discuss possible explanations and implications of this in view of the accident/conflict data. Here follows a systematic presentation of the different aspects that have been studied.

1. Accident data from ***separated*** roundabouts shows that the most common feature of cycle safety at the separated roundabout is the cycling direction; Regarding entering vehicles, most accidents occur with cyclists coming from the right, i.e., against the circulating direction. Accidents involving exiting vehicles are in most cases with cyclists coming from the left, i.e., also against the circulating direction. It seems as if entering and exiting vehicles entail more or less similar risks for cyclists. Earlier research shows that exiting vehicles are involved in more accidents than entering vehicles. The reason why they seem to be more alike here may be some factors that point in different directions. Drivers of exiting vehicles face a higher degree of complexity after interacting with other motor vehicles in the roundabout. At the same time results from our roundabouts in Lund indicate lower speeds when exiting, thus incurring lower risks. A third factor may be that drivers and cyclists feel more comfortable with the interaction when entering a roundabout, because there are still no other competing demands. This might lead to less attentiveness from both motorists and cyclists.

At ***integrated*** roundabouts, accident data shows that motor vehicles entering the roundabout (Int 1), i.e., vehicles not yielding to circulating cyclists, represent by far the biggest problem. Conflict data confirms this result; 4 out of 10 conflicts are Int 1. The remaining problems are linked to motorists and cyclists appearing in parallel situations. Int 3 – where motorists leave the roundabout, while the cyclist is still on it, is the largest problem in the parallel situations in terms of accidents. Since the numbers used here are small, further investigation is required.

2. There is a great difference in complexity between the roundabouts. Both the number of conflict types and the number of interaction types are higher in the integrated roundabout than in the separated one. The interactions at the bicycle crossing in the separated roundabout are similar to those at other types of intersections, hence both motorists and cyclists should be used to this kind of interaction. Still, both cyclists and motorists seem to be much more uncertain at this kind of crossing. Interactions at the integrated roundabout, however, are supposed to work like those between two motor vehicles in any roundabout. Nevertheless the situation turns out to be substantially different, as the cyclist can enter in parallel with the motorist, the motorist can catch up with and drive in parallel with the cyclist on the roundabout and the traffic rules often seem to be ignored. The consequences for accidents are further discussed in point no.5.

3. Interacting cyclists and motorists continue with unadjusted speed to a larger extent in the integrated roundabout than in the separated roundabout. In addition to better mobility in the integrated roundabout, it probably also leads to less attentiveness towards each other and higher risk for cyclists. The cyclist and motorist meet perpendicularly in the separated roundabout and hence it is quite obvious for the road

users involved that one of them has to yield. That does not mean, however, that the yielding rules are clear for either vehicle drivers or cyclists. In most of the interaction situations (all but Int1 & Int2) in the integrated roundabout the yielding rules are not obvious either. A cyclist may – as long as he/she does not perceive any high risks – easily enter the roundabout at the same time as a car, just by swerving slightly and continuing in parallel. The fact that no action (except for a slight swerving) is needed in most bicycle-motor vehicle interactions in the integrated roundabout could lead to motorists not noticing the cyclists but primarily looking for motor vehicles. The low level of conspicuity of cyclists is generally a problem in interactions with motor vehicles. It is also well known that drivers develop visual scanning strategies that enable them to scan the most important directions to avoid collisions with other motor vehicles (*Moray, 1990, Hills, 1980*). However, the development of scanning strategies also includes masking less important information such as cyclists. A motorist that is used to the presence of cyclists concentrates attention towards other motorists or things more likely to be important.

4. In the integrated roundabout 1/3 of the interactions are of Int1 or Int2 types. Int 1 (circulating cyclist, entering motorist) seems to be by far the most risky situation for cyclists in integrated roundabouts according to the accident statistics and conflict data. "The opposite situation", Int2 (circulating motorist, entering cyclist) is not at all as common in our accident statistics. The conflict data suggests that this might be because these conflict situations can more often end up in parallel driving if the road users (specifically the cyclists) adjust direction slightly. The consequence might be that Int 2 accidents will be avoided but that the cyclists will end up in parallel on the roundabout which may then lead to "parallel conflicts and accidents", i.e., Int 3, 4, 5 and 6.

The majority of the interactions (2/3) in the integrated roundabout are parallel situations. When motorists come up parallel with the cyclist, the latter faces problems finding a position where he can still be in control of the event. Of the parallel situations, Int3 (exiting motorist and circulating cyclist) seems, according to our studies, to be the most risky situation, which is in accordance with earlier research stating that Int 3 is the second most common accident situation at roundabouts (*Hels & Orozova-Bekkevold, 2007, Jørgensen & Jørgensen, 2002*). In 6 of the 23 Int 3-situations, for which we have noted the behaviour, the motorist continued at unchanged speed and the cyclist had to yield. It seems reasonable that this type represents the biggest problem of the parallel situations. The cyclists are probably not very prepared for the direction change of the motorist, and the motorist has great difficulties in observing the cyclist riding alongside, especially when the cyclist appears from behind. Cyclists also perceive Int3-situations to be the most dangerous (*Möller et al., 2000*). The risk with parallel driving, apart from ending up in Int3-situations, is that the cyclist may be squeezed by the motorist. (One of the reasons for roundabouts being safer than other intersections for motor vehicles is the small conflicting angle. However, cyclists, as vulnerable road users, may be just as badly hurt when the collision angle is small). The main problem with integrated roundabout used in this study is that the carriageway is wide enough to allow a cyclist and a car, or even a heavy vehicle, to drive side by side. An added problem is that

motorists do not seem to recognize parallel driving as a risky behaviour. About half of the motorists catching up with a cyclist stay behind when entering and circulating in the roundabout. When exiting, however, only 32% stay behind the cyclist. The most dangerous parallel situation is probably when the vehicle involved is heavy (bus or lorry); 5 of the 8 heavy vehicles that caught up with cyclists circulated in parallel. Even though the speeds are often low, the consequences of a cyclist being squeezed by a lorry or bus may be fatal. The cyclist, for obvious reasons, is cautious about moving up parallel with the motor vehicle; in 70% of the interactions where the cyclist caught up with a motor vehicle the cyclist chose to stay behind rather than in parallel. A conclusion is that parallel driving is not recommendable in this type of roundabout. It would, therefore, be interesting to find out whether there are any possibilities of creating a new design that will prevent parallel driving in the integrated roundabout, without creating new types of hazards. Expanded experiments to study this systematically are therefore warranted.

5. The yielding rules are interpreted – and executed – differently. In the separated roundabout both road users are obliged to yield. Motorists yielded to cyclists in 68% of the interactions, and in 30% of the interactions both road users adjusted their speeds or directions i.e., indicating they were not sure of the following behaviour of the other road user. In the integrated roundabout only one of the interacting road users has to yield, e.g. the entering vehicle has to yield to the circulating one. In Int1 situations (entering motorist, circulating cyclist) 96% of the motorists yielded and in Int2 situations (entering cyclist and circulating motorist) 86% of the cyclists yielded. Compared to the separated roundabout, only in 8% of the interactions did both the cyclist and the motorist adjust their speeds or directions in the integrated roundabout. The safety consequences of the difference in yielding behaviour between the two types are difficult to predict. Generally, one can say that the more ambiguous yielding rules in the separated roundabout not only lead to higher attentiveness and lower mobility for cyclists, but presumably also to lower risk. In the integrated roundabout a very high proportion of both motorists and cyclists yield according to the rules. Hence most motorists and cyclists rely on the other to yield as required, which causes safety problems once somebody violates the yielding rules, as the other part in such a case is not prepared to handle this situation (*Svensson, 1998*).

6. Bicycle crossings at roundabouts that allow cycling against the circulating direction seem to create a safety problem. Our results confirm those of *Räsänen & Summala, 2000*; i.e., motorists entering a roundabout yield more to cyclists coming from the left (circulating direction). Those cyclists are more conspicuous because motorists primarily look for other motorists, and therefore look in the direction from where the cyclist is coming. Consequently there seem to be higher risks for cyclists coming from the other direction. Several studies of different four-way intersections have also shown that cycling on the left side of the road on bi-directional cycle paths runs a higher risk of being involved in accidents (*Summala et al., 1996, Linderholm, 1992*). In the case of roundabouts only entering motorists' behaviour has been studied before, as mentioned above, showing that entering motorists yield more to cyclists riding in a circulating direction *Räsänen & Summala, 2000*. We also show that exiting motorists yield more to cyclists from the left, i.e., cyclists riding *against* the circulating

direction. Exiting motorists seem consequently to be more aware of cyclists going in the "wrong" direction. The cyclists from the left might also be more conspicuous to the driver because they have to pass one lane before the conflict point. However, even though interactional data seem to produce more favourable results for cyclists riding against the circulating direction, our accident data seems to verify the results from ordinary four-way intersections, namely that cyclists riding against the circulating direction are more at risk. Additional studies focusing more thoroughly on motorists' attention and cyclists' risk should therefore be performed to expand and explain these results.

Cycling in the wrong direction in the integrated roundabout should in practice have been impossible, but a small number (3%) of the cyclists still managed to ride in the wrong direction. In addition, 5% rode completely or partly on the pavement/zebra crossing at the integrated roundabout. The safety implications of this are unknown to us.

7. Our study employs automated video analysis as a detection method, which decreased the amount of video recordings to be watched manually by nearly 2/3 (from 90 to 35 hours). However, the quality of the detections varied a lot between the locations. One of the most obvious factors affecting quality is the position of the camera in relation to the road users. The problem of finding a good place for mounting the camera was evident early in the project. Many issues are to be considered here. The location has to provide a good view over the scene, have a power supply, and be easily accessible for installation but not for sabotage, etc. At both sites, the alternatives were quite limited, and the camera locations less than optimal. At separated roundabout the camera was too far from the studied bicycle crossings. The consequence was that cyclists had a small pixel size and were therefore often lost by tracking algorithms. The analysis of the "misses" shows that many of the missed cyclists had been correctly tracked for a while, but just near the detection gates there was some split in the trajectory, which made them undetectable. As it was difficult to orient a camera so that all the studied directions had the best view, and it was necessary to compromise. Approach C (gates I-1 and I-2) was given the highest priority since it had higher bicycle flows, which explains the lower detection rates at the other studied approach D (gates I-3 and I-4).

At integrated roundabout, on the contrary, the camera was too close to the site. Besides the obvious problem of the sight limitations (only half the roundabout was seen), the vehicles that came close to the camera were often detected as a group of cyclists, which produced a high number of false positives (90% at gate II-2 and 65% at gate II-3). Again, the camera was oriented to get the best view of gate II-1, where the detection and false positive rates were quite satisfactory.

An important question here is the reliability of video analysis as a detection method. If the misses occur just by chance, even very low detection rates can be compensated for by extending the observation period until the total number of detections is high enough to be able to draw significant conclusions. As video analysis does not require an operator to monitor the process, this is not a serious problem. However, if the misses occur due to some conditions related to the studied phenomenon (in this case

the safety of cyclists), the detections become systematically biased and the results might be misleading. For example, it was noticed that more cyclists were not detected when traffic became dense, and that they were more often occluded by other vehicles. If more conflicts (or accidents) occur in dense traffic, there might be more conflicts missed during high traffic hours than during low traffic hours. In such a case it would be incorrect to make a direct comparison of the conflict numbers of low and high traffic periods or two sites differing in traffic load. Other factors that affected the detection quality were the atmospheric conditions (rain, fog), light conditions and sun glare in the morning and evening hours and in some cases the cyclists' clothing (in some conditions the colour of their clothing was close to the colour of asphalt, and these cyclists were not detected at all). Further quality tests performed for different conditions are necessary to ensure that the system provides reliable results.

The number of false positives is unacceptably high. Partly, this can be explained by poor camera locations, but as it does not appear that finding good locations in the future will be easier, better algorithms for distinguishing between the road users of different types are necessary.

In addition to the obvious need for larger studies in order to link interactional behaviour with accidents and injuries in a more comprehensive way, it is important to approach the road users themselves. One reason is to better understand both motorists' and cyclists' comprehension of the different situations and to find the underlying reasons for their behaviour. Finally it is also important to ascertain what risks they perceive and to what extent high risks represent a threat to the general well-being of these road users.

6. CONCLUSIONS

The long discussion on which roundabout design is the safest for cyclists has not yet resulted in a definite answer. However, our results support earlier findings that roundabouts with separated bicycle crossings seem to be the safest. More important still is that risky situations have been identified and the road users' behaviour in those situations has been described, which will provide the background for designing more comprehensive studies of those situations. The integrated roundabout is more complex with a higher number of conflict and interaction types. Moreover, the yielding situation is clearer in the integrated roundabout, leading to a higher yielding rate but also to a greater trust in the other road user's willingness to yield as required. Hence the motorist and the cyclist are less prepared to act when either fails to yield.

The most dangerous situations in the integrated roundabout seem to be when the motorist enters while the cyclist is circulating, and when the motorist exits while they are circulating in parallel. In the separated roundabout the situations with the lowest yielding rate to cyclists occur when the motor vehicles exit the roundabout at the same times as cyclists are riding in a circulating direction and hence coming from the right. Still, most of the accidents in separated roundabouts take place when motorists enter or exit the roundabouts while cyclists are moving against the circulating direction. More emphasis should be put on these kinds of situations in order to understand the underlying safety implications.

Video analysis performs fairly well in studies where simple detections are necessary. It can "condense" the video material by selecting more relevant situations to look through, which enables the observation periods to be increased, thus providing more significant power to the results. The quality of the detection, however, still has to be improved with regard to both the detection and false positive rates. Further reliability tests are necessary to ensure that there is no systematic bias in missed detections.

7. REFERENCES

Ardö, H. (2009). Multi-target tracking using on-line viterbi optimisation and stochastic modelling. Doctoral thesis, Lund University, Faculty of Engineering LTH, Centre for Mathematical Science.

Aultman-Hall, L., F. L. Hall (1998). Ottawa-Carleton commuter cyclist on- and off-road incident rates. Accident Analysis and Prevention 30 (1), pp. 29-43.

Brüde, U., J. Larsson (1999). Trafiksäkerhet i cirkulationsplatser för cyklister och fotgängare (*in Swedish*) Traffic safety for pedestrians and cyclists at roundabouts. Statens väg- och transportforskningsinstitut (VTI). VTI meddelande 864.

Daniels, S., E. Nuyts, G. Wets (2008). The effects of roundabouts on traffic safety for bicyclists: an observational study. Accident Analysis and Prevention 40, pp. 518-526.

Elvik, R., T. Vaa (2004). The handbook of road safety measures. Elsevier.

Hels, T., I. Orozova-Bekkevold (2007). The effect of roundabout design features on cyclist accident rate. Accident Analysis and Prevention 39, pp. 300-307.

Herslund, M.-B., N. O. Jørgensen (2003). Looked-but-failed-to-see-errors in traffic. Accident Analysis and Prevention 35, pp. 885-891.

Hills, B. L. (1980). Vision, visibility and perception in driving. Perception 9, pp. 183-216.

Hydén, C. (1987). The development of a method for traffic safety evaluation: the Swedish traffic conflict technique. Doctoral thesis, Lund University, Department of Traffic Planning and Engineering.

Hydén, C. (1987). The development of a method for traffic safety evaluation: the Swedish traffic conflict technique. Department of Traffic Planning and Engineering, Lund University. Bulletin 70.

Hydén, C., A. Várhelyi (2000). The effects on safety, time consumption and environment of large scale use of roundabouts in an urban area: a case study. Accident Analysis and Prevention 32, pp. 11-23.

Jørgensen, E., N. O. Jørgensen (2002). Trafiksikkerhed i rundkørsler i Danmark - en analyse af uheld i danske rundkørsler i årene 1991-1996 (*in Danish*). Vejdirektoratet. Rapport 235.

Laureshyn, A., H. Ardö (2006). Automated video analysis as a tool for analysing road user behaviour. ITS World Congress, London, UK.

Laureshyn, A., H. Ardö, T. Jonsson, Å. Svensson (2009). Application of automated video analysis for behavioural studies: concept and experience. IET Intelligent Transport Systems 3 (3), pp. 345-357.

Linderholm, L. (1992). Utvädering av trafiktekniska åtgärders säkerhetseffekter - metodutveckling med tillämpning på utformningsdetaljer för cyklister i signalreglerade konsrningar (*in Swedish*). Lund University, Department of traffic planning and engineering.

Moray, N. (1990). Designing for transportation safety in the light of perception, attention and mental models. Ergonomics 33, pp. 1201-1213.

Möller, E., G. Grieszbach, B. Schack, H. Witte, P. Maurizio (2000). Statistical properties and control algorithms of recursive quantile estimators. Biometrical journal 42, pp. 729-746.

Räsänen, M., H. Summala (1998). Attention and expectation problems in bicycle-car collisions: an in-depth study. Accident Analysis and Prevention 5, pp. 657-666.

Räsänen, M., H. Summala (2000). Car drivers' adjustment to cyclist at roundabouts. Transportation Human Factors 2, pp. 1-17.

Schoon, C., J. van Minnen (1994). The safety of roundabouts in the Netherlands. Traffic Engineering and Control 03 (35(3)), pp. 142-143, 145-148.

Summala, H., E. Pasanen, M. Räsänen, J. Sievänen (1996). Bicycle accidents and drivers' visual search at left and right turns. Accident Analysis and Prevention 28, pp. 147-153.

Svensson, Å. (1998). A method for analysing the traffic process in a safety perspective. Doctoral thesis, University of Lund, Lund Institute of Technology, Department of Traffic Planning and Engineering.
TRF (2004) Trafikförordning. 6 kap 6 §, 3 kap 61 §, 3 kap 22 §. www.notisum.se.