Automatic detection of mobile interactions in a signalized intersection
Saunier N., Midenet S., Grumbach A.
Proceedings of the 16th ICTCT Workshop, Soesterberg, The Netherlands, 2003

AUTOMATIC DETECTION OF MOBILE INTERACTIONS IN A SIGNALIZED INTERSECTION[1]

Nicolas Saunier*[#], Sophie Midenet*, Alain Grumbach[#]

*French National Institute for Transport and Safety Research (INRETS)
Transport Network and Advanced Software Engineering laboratory (GRETIA)
2 avenue du Général Malleret Joinville, 94114 Arcueil, France
Phone: +33(0)147407292, Fax: +33(0)145475606

[#]Telecom Paris (ENST)
Network and Computer Science Department (INFRES)
46 rue Barrault, 75013 Paris, France
Phone: +33(0)145817532, Fax: +33(0)145813119

Email: {saunier, grumbach}@enst.fr, midenet@inrets.fr

ABSTRACT

This work deals with the consequences of traffic light control strategies in signalized intersections on the behavior of users, particularly in terms of risk. Traffic light control consists in adapting traffic lights to the variable demand of traffic in order to improve the fluidity and the safety. An adaptive real-time control strategy developed at INRETS has been tested and gave good results according to fluidity criteria. Traffic in a complex real intersection was recorded over a period of eight months, during which time four different control strategies were applied. The resulting database allows us to make a comparative analysis of the effects of the different strategies. It is based on an experimental automatic observation device supplied by video sensors, which gives us dynamic measurements of the spatial occupation of the intersection.

We examined traffic events relevant to safety. As no accident occurred during the experiment, our interest turned to traffic conflicts and their observation methods. Following the Swedish traffic conflict technique and Åse Svensson's work (Å. Svensson, 1998), we study traffic events relevant to safety. We apply her representation of the number of events for each degree of severity, which allows more complete and subtle diagnosis. Within this framework, we propose a categorization of traffic events in a signalized intersection, with respect to the features of the occupation measurements. By using pattern recognition methods to deal with these measurements, this approach treats very large databases in a systematic way. We define three categories of interactions, according to various speeds and positions of conflict protagonists. We explain the detection methods, and how we plan to evaluate the severity.

1The research reported in this paper is part of a PhD thesis, supported by a region Ile de France grant.

1. INTRODUCTION

When we think of signalized intersections, we think of traffic lights, and their cycles. If we interviewed road users about their feeling towards them, many would answer that they waste a lot of time waiting. We can indirectly have an insight into this feeling by counting the rate of run-red offences in urban areas (S. Midenet, 1998). If people could control the traffic light, they would probably enhance the fluidity and the capacity. But such a change can also have an influence on safety. A traffic more fluid usually means higher speed, which in turn means higher risk of accident, and more serious potential injuries. This illustrates the fact that traffic light control has an influence on the behavior of road users, and their safety.

The Problem

The purpose of our work is the comparison of traffic light control strategies, and their influence on the behavior and the safety of road users. We evaluate this influence by automatically detecting mobile interactions in a signalized intersection. This work is part of a research project dealing with traffic management on signalized junctions, based on video sensors and a real-time traffic light control strategy. An experiment was run over a period of eight months on a signalized intersection near Paris to compare this strategy to references, yielding a large database of traffic scenes at our disposal for this work.

The paper is composed as follows: section 2 presents our approach and the reason for focusing our study on interactions occurring in the conflict zone. Three categories of interactions are singled out in section 3. Section 4 describes the surface data and the concrete definitions of the categories of interaction and of our severity indicators. In section 5, we develop the task and an architecture to solve it. Section 6 presents encouraging results. Section 7 concludes, and perspectives are discussed.

2. OUR APPROACH

Intersections are critical places with respect to the safety of road users. In particular, the most crucial phase is the crossing of the conflict zone. The conflict zone is the place where the mobiles coming from different streams cross, and is thus potentially dangerous. Traffic lights are supposed to prevent cross streams from entering the conflict zone simultaneously. The dangerous situations that however occur there may have a link to the way traffic lights are controled as described in the introduction. For this reason and considering our purpose of studying the influence of traffic light control on safety, we focus our work on traffic events occurring in the conflict zone, evaluating the recurrent conditions of mobiles crossing intersections.

We can draw a parallel between our approach and the studies of the drivers' decision-making at signalized intersections during signal change interval, and particularly when optimizing the yellow or green timing, with or without a mobile-actuated strategy (R. van der Horst et al., 1986) (R. van der Horst, 1988) (D. Mahalel et al., 1985) (R. R. Retting et al., 2002). In the same way as these studies deal with the influence of traffic light control in a certain temporal window (during a phase change), our approach deals with the same problem in a certain spatial window, that of the conflict zone. Besides, these works show that minor changes in traffic light control do have an influence on the behavior of road users, and their safety, the most obvious being the evolution of the rate of run-red offences.

What kind of traffic events can give us an insight into the safety situation of an intersection ? Most safety studies are based on the observation of accidents, traffic conflicts, or both. There are no accidents in the database, and there are theoretical problems about their use (G. R. Brown, 1994). On the other hand, traffic conflicts are recurrent events, more frequent than accidents. The Traffic Conflict Literature (G. R. Brown, 1994) (C. Hyden, 1987) (N. Muhlrad, 1988) (T. Sayed et al., 1999), and especially Åse Svensson's work (Å. Svensson, 1998) supply us with a framework for the description of interactions between road users, from normal situations to accidents. In this framework, we want to detect relevant traffic events automatically to statistically compare different traffic light strategies. All *interactions* between road users occurring in the conflict zone, with or without a collision course are suitable.

Not only will the interactions be detected and counted, but we will also try to quantify their severity, defined as the distance between the interaction and the potential accident. We are familiar with the classical indicators, simple indicators like speed and distance, and more elaborate indicators like Time To Collision (TTC) or Post-Encroachment Time for interactions with a collision course (see (R. van der Horst, 1990) for their definition and other indicators). Our indicators however are chosen in function of the features of the data.

It is not easy to interpret traffic events and their indicators as a measure of safety. For example the link between accidents and traffic conflicts is not established. We will use Åse Svensson's severity hierarchy, i.e. the representation of the number of events for each degree of severity. The interpretation of the shape of the representation allows a more complete and subtle safety diagnosis.

Safety studies comparing signalized intersections and non-signalized intersections exist. Åse Svensson (Å. Svensson, 1998) shows that hierarchy shapes in signalized and non-signalized intersections are different, i.e. at non-signalized intersections, the distribution of interactions is located towards higher severity values than at signalized intersections. However, as far as we know, the influence of different traffic light strategies on the safety of road users has never been compared.

3. A CATEGORIZATION OF INTERACTIONS

We call road users and their vehicles *mobile*, i.e. pedestrians or other road users with their vehicle. We're interested in interactions between all types of mobiles, but there are limitations on account of the characteristics of the data. Neither can we perceive the type of mobiles, nor can we distinguish what is happening in homogeneous groups of mobiles. As a consequence, our approach is relatively rough, based on zones, and the presence of mobiles in them, assuming that mobiles in storing zones must cross the conflict zone at one time or another. This assumption constitutes our extrapolation of trajectories to decide the existence of collision courses. We detect interactions, with or without collision courses, on the level of the zones.

Among interactions between mobiles occurring in conflict zones we singled out the following categories (see illustration 1), discriminated according to the speed and position of the protagonists in zones,

- *downstream category*: interactions with a collision course between a mobile or a group of mobiles entering the conflict zone and a mobile or a group of mobiles stopped in the conflict zone,
- *stationary cross traffic category*: interactions between a mobile or a group of mobiles in the conflict zone and a mobile or a group of mobiles stopped in the cross traffic[2] storing zone,
- *moving cross traffic category*: interactions with a collision course between a mobile or a group of mobiles entering the conflict zone and a mobile or a group of mobiles moving in the cross traffic[2] storing zone.

As far as the stopped cross traffic category is concerned, we consider that this situation is not neutral with respect to the risk of accident, as the mobile in the cross traffic storing zone could move off, and because we want to discriminate between this situation and a safe situation where the cross traffic conflict zone is empty.

2The cross traffic storing zone is the storing zone different from the zone the mobile or group of mobiles located in the conflict zone comes from.

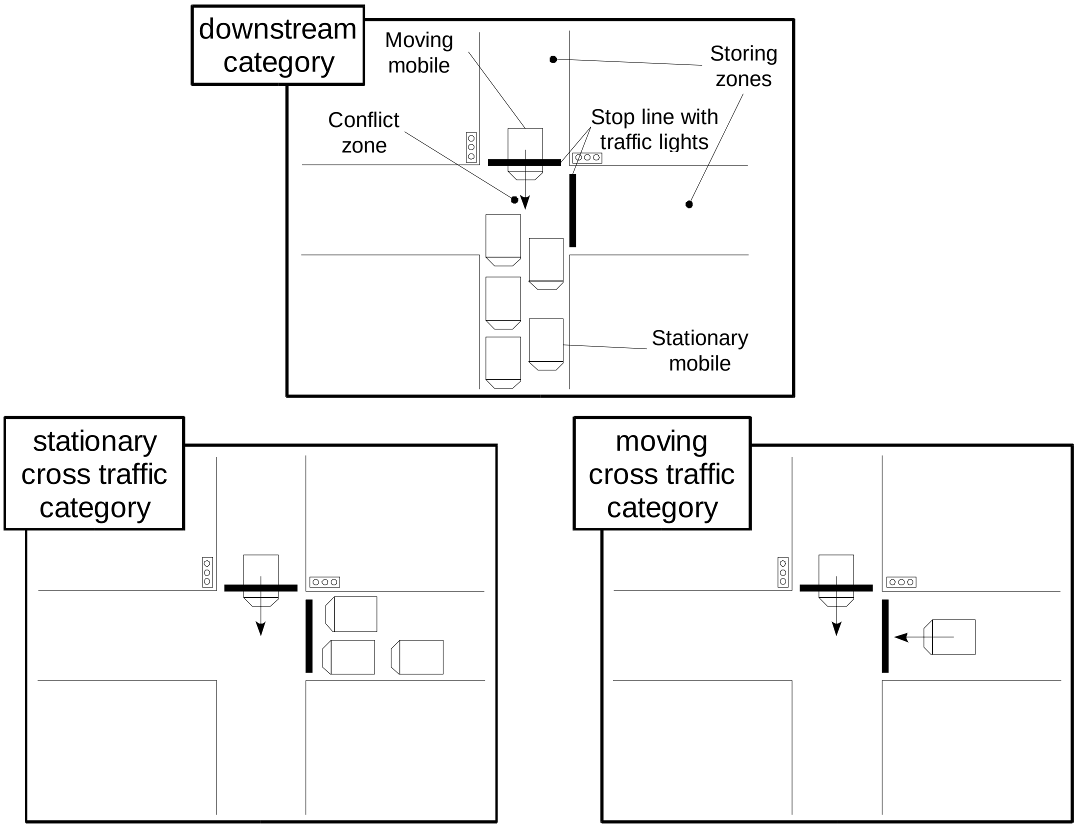

Illustration 1, The categories to be detected

This categorization doesn't cover all interactions between groups of mobiles in an intersection, even when restraining ourselves to interactions potentially related to traffic light control. Nonetheless, the chosen categories should yield interesting results as far as comparing traffic light control strategies is concerned.

4. SURFACE DATA FROM VIDEO SENSORS

Description of the intersection

The experimental site is a junction located in an urban area near Paris, in the Val de Marne department. Two roads with asymmetric traffic cross at this junction: a main road linking Paris to the A6 motorway and a secondary road connecting two Val de Marne cities. The junction is considered isolated as it is quite far from other signalized junctions. Fifteen traffic lights (among which seven for pedestrian crossing) control the junction. The junction is composed of four intersections of two one-way roads. Its inner part is composed of four conflict zones and four storing zones for the left-turning mobiles controlled by a traffic light group (see illustration 2). Eight cameras have been installed in the junction, covering the inner part and the approaches.

The data

The video images coming from the eight video cameras are processed by a robust tool, developed at INRETS, built to work twenty-four hours a day, in any weather, to provide two-dimensional discrete images of the occupancy of the junction. The data are used by several modules of the traffic management system, beginning with the traffic light control module. The measurements are much richer than what magnetic loops can offer, spatially and temporally. The research described in this paper concerns the interpretation of these measurements in order to detect and qualify traffic events related to safety. As far as we know, there exists no automated device dedicated to the detection of traffic conflicts, only semi-automated devices or projects (K. Odelid et al., 1993) (J. Andersson, 2000) (H. Veeraraghavan et al. 2002).

The basic information units processed several times per second from the video images are emptiness, presence of moving mobiles and presence of stationary mobiles. To have access to the dynamics of the mobiles, we combined this information to generate every second higher-level detection patterns (see illustration 3), which we call from now *images*.

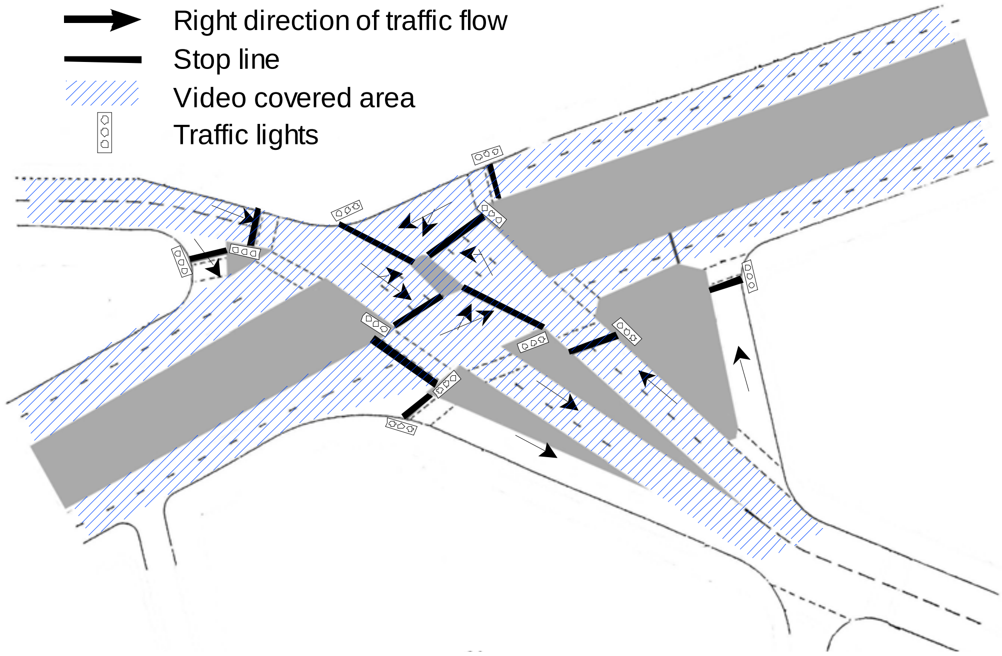

Illustration 2, The experimental junction

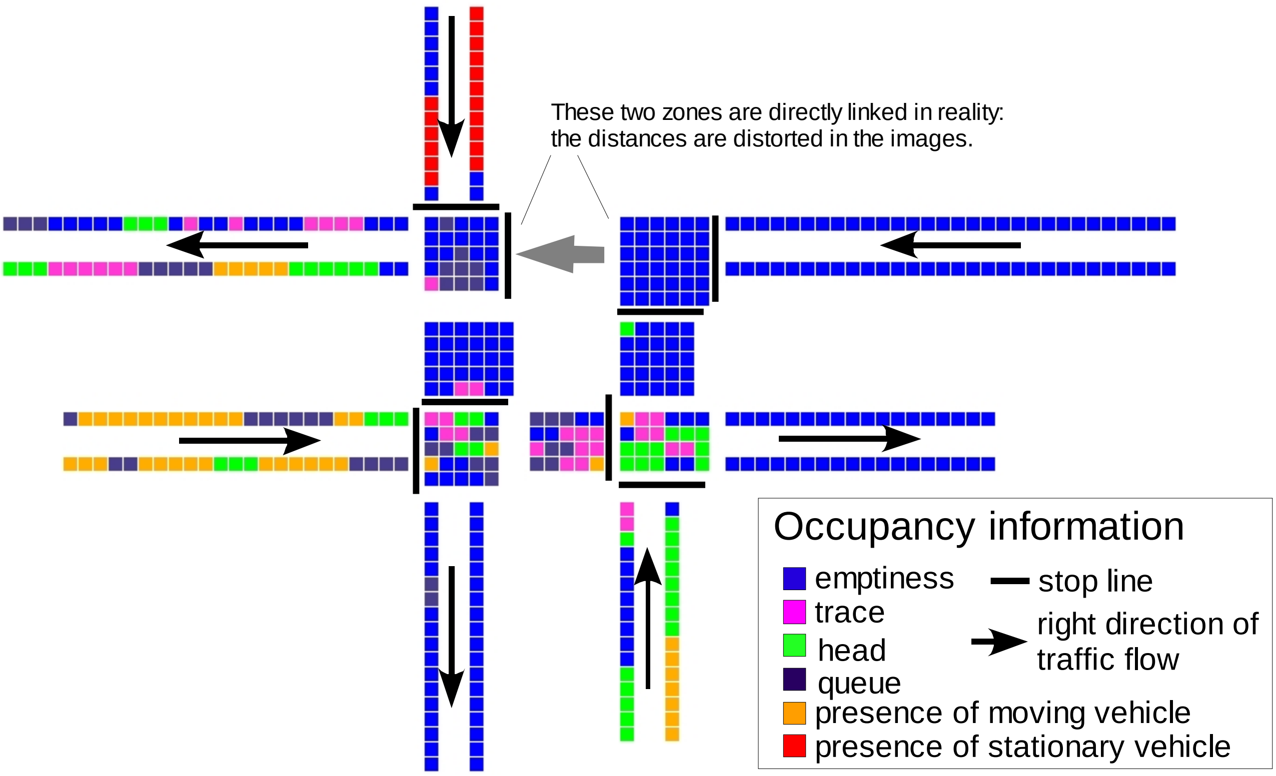

Illustration 3, An image example (The names of occupancy information are explicit, except for trace which means emptiness at times t-1 and t, and perceived presence in between, indicating a quick movement)

The information is bi-dimensional for the inner junction and linear per lane for the approaches to the junction. The information units indicate the presence of mobiles over a period of one second, say from t-1 to t, i.e. the presence at times t-1 and t, and the dynamics during the second since t-1 (see illustration 4). This enables us to handle the diversity of mobile movements in a robust way.

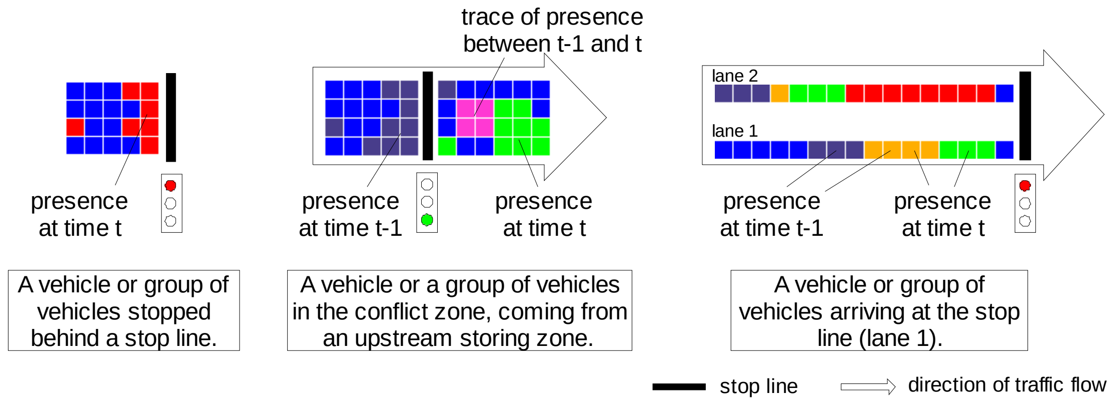

Illustration 4, Some clear typical patterns in the data

Interactions in the data

We cannot detect the type and the spatial limits of the mobiles, since the only basic information is absence and presence. Therefore, we can only detect groups of mobiles, and interactions between groups. We build connected sets of units of presence of stationary mobiles on the one hand and connected sets of units of presence of moving mobiles on the other hand, which indicate the presence and dynamics of the detected groups of mobiles. These connected sets of units are called *blobs*, and are identified as part of the interaction, as the traces of the protagonists in the images.

Detecting presence and movement in the different zones, and assuming that the mobiles follow the right direction of traffic flow, we conclude the interaction exists or not (see examples on illustration 5).

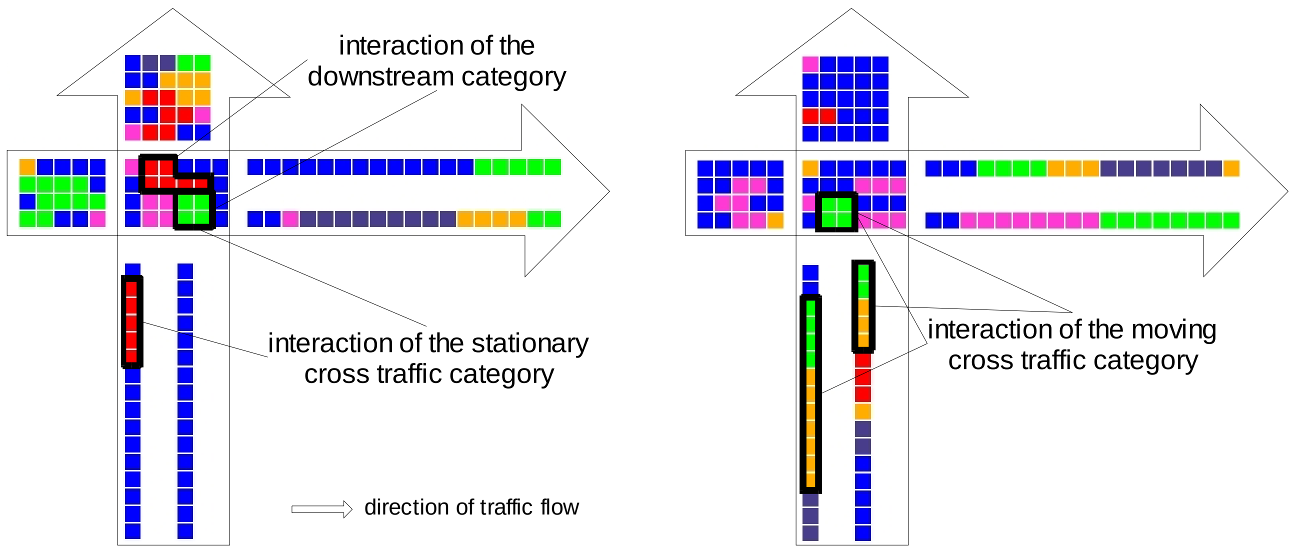

Illustration 5, Examples of interactions in the data

Severity indicators

The information contained in the images are the presence and dynamics of groups of mobiles, hence we can have an idea of distances between protagonists and their speed. For interactions with a collision course, the evasive action cannot be detected in most cases, on account of the definition of the data. As we can only estimate indicators like distances and speed, we reject their combination into indicators like the Time To Collision which would worsen the precision. Neither will we build a unique severity indicator both for this reason and because of the difficult process of validation of a new severity indicator. We will rather measure meaningful basic indicators like the speed differential and the proximity to have an indication of the severity. The higher the speed differential, or the closer the protagonists, the more severe the interaction. But we are not able to say if an interaction with two mobiles very close but moving slowly is more or less severe than another interaction with two mobiles moving very fast, but far from each other. Our indicators are related to the severity, defined as the distance of the interaction to the accident.

On account of the spatio-temporal definition of the data, there will be only a few discrete values for the indicators, like minimum, maximum and medium. We define two indicators, standing for the severity information contained in the data,

- the *extrapolated proximity*, defined as the minimal extrapolated distance between the protagonists,
- the *speed differential* between the protagonists, defined as the length of the difference of their speed-vectors at the time of detection.

The data and the indicators are quite rough, compared with what human observers can perceive, but we are not trying to implement a Traffic Conflict Technique in an automatic detection device. We want to detect the categories of interactions described above in a systematic way, which then allows us to process a large database like ours automatically.

5. OUR DEVELOPMENT

The task and the architecture

We want to detect and qualify automatically the interactions of the chosen categories present in our database of images, recorded under different strategies. One could think of trying to derive the kinematics of mobiles from the data, but the fact that the mobiles are not perceived separately makes this approach difficult. We prefer working on images and try to make the most of them separately.

Then, for each image at time t, we will detect the interactions of the three categories, downstream, stopped cross traffic and moving cross traffic, that happened over the preceding second or are happening at time t, and qualify them according to the chosen indicators.

Therefore a modular architecture was developed to deal with the data (see illustration 6),

- a first module, detecting the interactions and their categories in an image, using a rule-based method,
- a second module, estimating the severity indicators of each interaction detected by the first module, using an explicit computation for the extrapolated proximity, and supervised learning for the speed differential.

Explanation of the methods and evolution of the task

Supervised learning is a set of learning methods used to teach a model, like artificial neural networks, a mapping between inputs and outputs. Once learned on a labeled set, the model is used on unlabeled data. This kind of method can generalize the target concept from labeled data, and is therefore generic, as opposed to an explicit evaluation. The images are complex and estimating the speed differential needs to take into account information spread over the whole image, as opposed to the extrapolated proximity which can be explicitly computed as a distance between interacting blobs in the images.

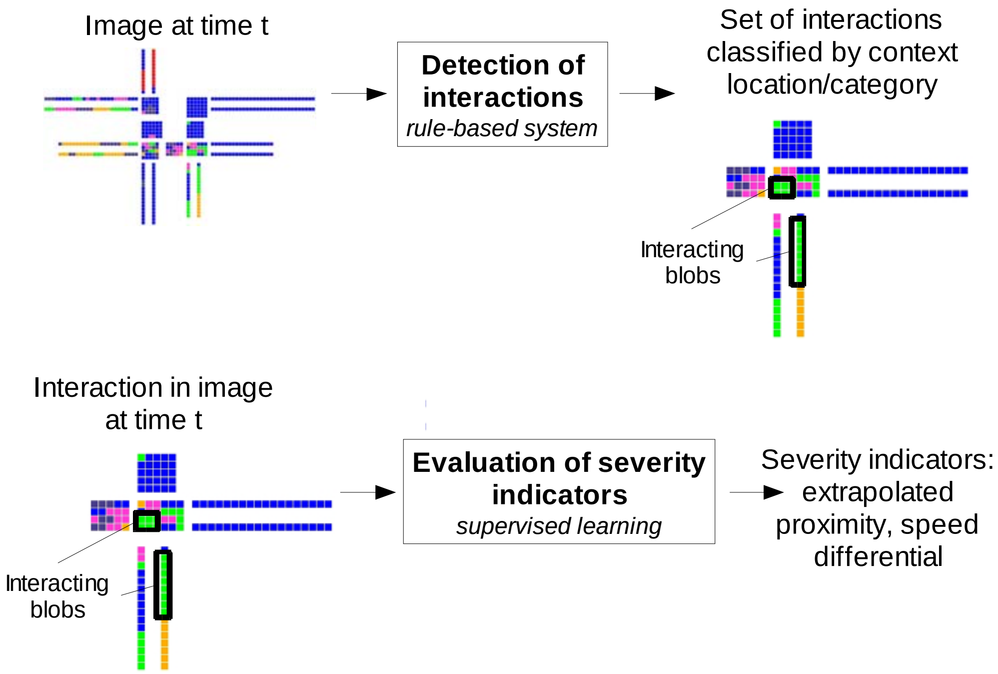

Illustration 6, The modular architecture of the system

Consequently, the system evaluating the speed differential must be supplied with labeled data, i.e. interactions detected in images, with the corresponding value of speed differential, among a few discrete values. A human expert must do the labeling, a rather complicated task. As described in the literature (J. Diez et al. 2002) (W. W. Cohen et al., 1999), we think that it is an easier task to compare the elements to each other in order to build the scale and levels of the indicator. Since the zones of the intersection are filmed by different video cameras, the distances and speeds, the indicators should not be compared from one zone to another. Therefore, we need to distinguish the interactions by the name of the conflict zone in which they are occurring, for the learning of the speed differential, and the comparison of interactions when interpreting the results, i.e. the distribution of the interactions according to each indicators.

We call *context* of an interaction these two attributes, its location (the conflict zone) and its category. Since we compare the speed differential of interactions separately in their contexts before learning, the learnt scales of this severity indicator may be different from one context to another.

Thus, the first module has to detect in every image the interactions of the three chosen categories, and their context attributes, category and location. We identified a few explicit rules for the mapping from configurations of positions of blobs to the categories. The interacting blobs are memorized to be used by the second module. This second module evaluates the indicators, extrapolated proximity and speed differential, for every detected interaction. The extrapolated proximity is directly computed as a transversal distance between interacting blobs. The speed differential is estimated by a model, adjusted in a prior learning phase with labeled interactions.

We also observed that more than one interaction can occur at the same time in the same conflict zone. In this case, if the input of the module is only the image, there is an ambiguity in the output. The images must be processed to specify which interaction the module is supposed to evaluate. The interacting blobs are memorized by the first module for that purpose.

How to focus on interactions ?

The interacting blobs indicate the parts of the images where the traces of the interacting groups of mobiles are, hence where the information necessary for the computation of the severity indicators is. We call *focusing* this problem of weighing the relative usefulness of the parts of the input. Three possible solutions are offered,

- take out the "unnecessary" parts of the images, i.e. everything except for the interacting blobs. This method is suitable for the extrapolated proximity, but not for the estimation of the speed differential, where the information can be scattered in the image outside of the strict blobs,
- add information in input to indicate where the system should look for the necessary information, which makes the outputs unambiguous,
- modify the learning algorithm of the model, taking differently into account the parts of the input in the learning process.

The extrapolated proximity is directly computed and follows the first solution. The second solution is experimented for the speed differential by adding to each unit of an image the information of whether it is part of an interacting blob or not. In this way, the model can learn where the most informative parts of the image are, and can still make the necessary correlation as for the rest of the image.

6. CURRENT RESULTS

The first module is developed, and has been tested qualitatively on a few minutes. The second module is being developed. Some prototypes were programmed, particularly for the testing of the focusing method described above (Nicolas Saunier et al., 2003). We use an artificial neural network, more precisely a multi-layer perceptron, as the model to learn the speed differential. The set of examples contained 623 interactions. Averaged over 10 generations of the learning and test sets, with 5 trials per set, the ratio of good prediction in generalization is 88 % (with 3 points as standard deviation). These results are encouraging.

7. CONCLUSION

This work, still in progress, needs to be pursued to achieve the initial goal, the comparison of the influence of different traffic light control strategies on the behavior and safety of road users. For that purpose, we isolated and categorized traffic events occurring in conflict zones, to be detected with pattern recognition methods, namely a rule-based module to detect interactions in data images, and supervised learning to learn a severity indicator, the speed differential. As a by-product, we present the focusing problem, some solutions, and current results.

We expect interesting results as to the transportation problem, and to provide new tools for the safety evaluation of traffic light control strategies and the monitoring of safety in intersections.

We plan to begin soon processing our database. We are thinking of using another important feature of the problem, the abundance of unlabeled data, i.e. the whole database, and techniques to take them into account to enhance the performance of the models. We also have to find a way to control the performance of the system all along its exploitation.

REFERENCES

The Swedish Traffic Conflict Technique, University of Lund, (http://www.tft.lth.se/rapporter/Conflict1.pdf).

J. Andersson, 2000, A new traffic analysis method based on image processing, University of Lund, Bulletin 193, 2000.

G. R. Brown, 1994, Traffic conflict for road user safety studies, Canadian Journal of Civil Engineering, Vol. 21 1994, pp. 1-15.

W. W. Cohen, R. E. Schapire, Y. Singer, 1999, Learning to Order Things, Journal of Artificial Intelligence Research, Vol. 10, 1999, pp. 243-270.

J. Diez, J.J. del Coz, O. Luaces, F. Goyache, J. Alonso, A. M. Peña, A. Bahamonde, 2002, Learning to assess from pair-wise comparisons, Proceedings of the 8th Iberoamerican Conference on Artificial Intelligence (IBERAMIA), 2002, Sevilla, Spain, Lecture Notes in Artificial Intelligence, LNAI 2527, pp. 481-490.

R. van der Horst, A. Wilmink, 1986, Drivers' decision-making at signalized intersections: an optimization of the yellow timing, Traffic Engineering Control, Vol. 27 1986, pp. 615-622.

R. van der Horst, 1988, Driver decision making at traffic signals, Transportation Research Record, Vol. 1172 1988, pp. 93-97.

R. van der Horst, 1990, A time-based analysis of road user behavior in normal and critical encounters, PhD Thesis, Delft University of Technology, 1990.

C. Hyden, 1987, The development of a method for traffic safety evaluation: The Swedish Traffic Conflicts Techniques, University of Lund, Bulletin 70, 1987.

D. Mahalel, D. Zaidel, T. Klein, 1985, Driver's decision process on termination of the green light, Accident Analysis & Prevention, Vol. 17 No. 5 1985, pp. 373-380.

S. Midenet, 1998, Self-organizing maps for spatial measure interpretation and traffic description in urban junctions, Nîmes 98, Special Session: Transportation and Applications of Neural Networks Technology (in French).

N. Muhlrad, 1988, Technique des Conflits de Trafic, Manuel de l'utilisateur, Synthèse INRETS n°11, 1988 (in French).

K. Odelid, Å. Svensson, 1993, A more automatic conflict technique with image processing, Proceedings ICTCT Workwhop Salzburg, 1993.

R. R. Retting, J. F. Chapline, A. F. Williams, 2002, Changes in crash risk following re-timing of traffic signal change intervals, Accident Analysis & Prevention, Vol. 34 Issue 2 2002, pp. 215-220.

Nicolas Saunier, Alain Grumbach, Sophie Midenet, 2003, Analyse symbolique-numérique de scènes de trafic par apprentissage focalisé, submitted to RFIA 2004 (in french).

T. Sayed, S. Zein, 1999, Traffic conflict standards for intersections, Transportation Planning and Technology, Vol. 22 1999, pp. 309-323.

Å. Svensson, 1998, A Method for Analyzing the Traffic Process in a Safety Perspective, PhD dissertation, University of Lund, 1998.

H. Veeraraghavan, O. Masoud, N. P. Papanikolopoulos, 2002, Managing suburban intersections through sensing, Technical Report, University of Minnesota, CTS 02-07, 2002.