Use of automated video analysis for the evaluation of bicycle movement and interaction

Heather Twaddle*[a], Tobias Schendzielorz[b], Oliver Fakler[c], Sasan Amini[a]
[a]Technische Universität München, Arcisstraße 21, 80333 Munich, Germany; [b]Heusch/Boesefeldt GmbH, Zenettistraße 34, 80337 Munich, Germany; [c]TRANSVER GmbH, Maximilianstraße 45, 80538 Munich, Germany

ABSTRACT

With the purpose of developing valid models of microscopic bicycle behavior, a large quantity of video data is collected at three busy urban intersections in Munich, Germany. Due to the volume of data, the manual processing of this data is infeasible and an automated or semi-automated analysis method must be implemented. An open source software, “Traffic Intelligence”, is used and extended to analyze the collected video data with regard to research questions concerning the tactical behavior of bicyclists. In a first step, the feature detection parameters, the tracking parameters and the object grouping parameters are calibrated, making it possible to accurately track and group the objects at intersections used by large volumes of motor vehicles, bicycles and pedestrians. The resulting parameters for the three intersections are presented. A methodology for the classification of road users as cars, bicycles or pedestrians is presented and evaluated. This is achieved by making hypotheses about which features belong to cars, or bicycles and pedestrians, and using grouping parameters specified for that road user group to cluster the features into objects. These objects are then classified based on their dynamic characteristics. A classification structure for the maneuvers of different road users is presented and future applications are discussed.

Keywords: Bicycle Behavior, Modeling Bicycle Traffic, Automated Video Analysis, Trajectory Data, Open Source

1. INTRODUCTION

Trends developing since the second half of the 20th century, including urbanization, suburbanization and the increasing dependence on personal motor vehicles, have led to severe traffic related problems in urban areas. To deal with congestion, air pollution, traffic accidents and noise, among other issues, city planners and traffic engineers are searching for new solutions to improve traffic safety and efficiency. A solution that has had considerable success in a number of cities, particularly in Europe, is the introduction of policies and infrastructure that support bicycle use. As a result, the modal split of bicycles has grown in the last decades and continues to rise today[1].

In order to be able to design and assess urban transportation infrastructure in accordance to the needs of all road users, including bicycles, knowledge concerning the characteristics and behavior of bicycle traffic is required. Currently, in comparison with the extensive body of knowledge regarding both the macroscopic characteristics of motorized traffic streams, including flow, density and average speed, and the microscopic behavior of drivers, such as car following and lane changing behavior, very little information exists concerning the flow of bicycle traffic and behavior of individual bicyclists. A large amount of observational data that quantifies the movement of bicyclists and their interactions with other road users in a number of situations must be collected and analyzed to create a similar body of knowledge for bicycle traffic.

Microscopic traffic simulation tools are commonly used to design and assess transportation infrastructure. As the proportion of bicycle traffic continues to rise in urban areas, the impact of bicycles on the overall flow of mixed traffic streams also increases, making the realistic inclusion of bicycle traffic in microscopic transportation simulation tools imperative for accurate assessments of transportation infrastructure. Because the behavior of bicyclists differs considerably from that of motorized vehicles, the behavior models used to depict road users in the simulation tools must be calibrated, adapted or extended to reflect the characteristics of bicycle traffic. In some cases, new behavioral models may need to be developed to capture unique aspects of bicyclists’ behavior.

In order to assess the state of the art in bicycle modeling and to identify areas for improvement, an extensive review of available bicycle models was carried out[2]. The results indicate that it is currently possible to model the operational behavior[3] of bicyclists, which is the short term (milliseconds and seconds) behavior that takes place on the subconscious level, including lateral and longitudinal spacing between road users, desired speed distributions, passing maneuvers and queuing behavior. It is necessary, however, to calibrate and validate these models using data from reality. There are fewer models currently available that depict the tactical behavior[3] of bicyclists, which is the conscious, short to midterm (seconds or minutes) behavior, including the selection of a trajectory across an intersection or cooperative behavior with other road users at intersections. Due to many factors, including their small size, their ability to ride on different infrastructure types and their possibility of switching between pushing and riding, bicyclists are much more flexible than motorized road users. This flexibility has a considerable influence on the effect of bicycle traffic on the overall flow of mixed traffic streams. The inclusion of flexible behavior in models takes place mainly on the tactical level, which enforces the need to develop and extend models on this level.

Additionally, an important aspect in the design of urban infrastructure is the protection of vulnerable road users including bicyclists. In Germany, while only 9% of the trips are made by bicycle, 21% of people fatally injured in urban traffic accidents in 2013 were bicyclists[1]. The design and construction of intersections and roadways that promote safe interactions between bicyclists and motorized road users is fundamental in decreasing the number of bicyclists that are injured or killed in traffic accidents. However, observational data concerning how road users interact in different situations must be collected as a first step in understanding how infrastructure design affects bicycle safety.

Finally, within the German research initiative UR:BAN (Urban Space: User oriented assistance systems and network management), advanced driver assistance systems (ADAS) and intelligent transportation systems (ITS) are being developed to improve traffic safety, increase efficiency and reduce harmful environmental effects in urban areas. The goal of a number of specific UR:BAN applications is to improve the safety of vulnerable road users including bicyclists. The accurate evaluation of these applications depends on the development of realistic models of bicycle behavior. The collection of observational data and the development of realistic bicycle models is one of the work packages within UR:BAN.

In order to develop a knowledge base concerning the movement and interactions of bicyclists, a large quantity of trajectory data from bicyclists, motor vehicles and pedestrians at a variety of intersections with different geometric and traffic characteristics is required. Within this paper, the methodology used in the UR:BAN work package for extracting road user trajectories using an open source software "Traffic Intelligence"[4], classifying the road users as bicycles, pedestrians or motor vehicles, and merging disconnected trajectories is presented. Video data from three intersections in Munich, Germany is used to test and evaluate the methodology. The resulting database contains the classification, position and speed of all road users throughout their path across the intersection, which can be used to extend knowledge concerning many facets of bicyclists' behavior.

2. METHODOLOGY

During the summer of 2013, video data was collected at three intersections in Munich, Germany for three days each. A methodology for extracting trajectory data from the video data was developed and has been tested using sample video data from each of the three test intersections.

Video data was selected as a medium to collect and extract trajectory data for a number of reasons. First, it is possible to gather detailed information about the situation, including the trajectory and type of all road users, the traffic situation, the current weather conditions and other unique circumstances. Second, a video camera can be easily mounted and dismounted, enabling the collection of data at many different intersections. Finally, the video camera system used was inexpensive and allowed for the collection of a large amount of data within the allotted budget. Despite these advantages, the processing of video data presents a multitude of challenges. Due to the quantity of required data, the manual processing of video data is infeasible and an automated or semi-automated video analysis tool must be used. The software "Traffic Intelligence"[4], which is an open source software developed by Nicolas Saunier and his team at the Polytechnique Montreal in Canada, was selected because of its accessibility and the potential to extend the software to analyze particular components of the UR:BAN research questions in an automated way.

As a first step in the automated video analysis, an optimization of the feature detection and object grouping parameters of “Traffic Intelligence” is carried out. A method developed for this project for separating the tracked features of motorized vehicle traffic from those of bicycles and pedestrians and grouping the objects separately is then presented and evaluated. Finally, a newly developed automated method for combining the disconnected trajectories of road users who start and stop or pass under or behind an obstructing object, such as an overhead cable or a traffic light pole, while crossing the intersection is described and evaluated. The developed methods are tested on a sample of video data from the three test intersections in Munich and the advantages and disadvantages as well as potential for improvement are discussed.

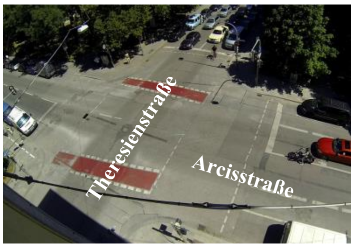
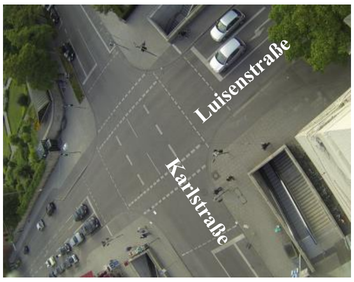
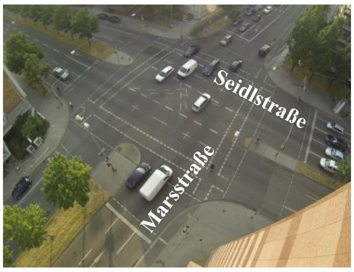
2.1 Data Collection

Munich was selected for data collection because of the relatively high modal split of bicycle traffic as well as the wide variety of bicycle infrastructure used to carry bicycle traffic. These aspects make it possible to deduct conclusions about bicycle behavior in many situations. In the last decades, the city has made a concerted effort to increase the proportion of trips that are made by bicycle[5]. Bicycle infrastructure, including bicycle lanes, both within and separated from the road way have been built, signalization for bicycle traffic has been implemented, bicycle parking in the city has been provided and a campaign continues to encourage people to bicycle. As a result, in 2011, 17.4% of all trips in the city were made by bicycle[6]. There are 13 types of infrastructure used in Munich to carry bicycle traffic, which can be divided into four main types, roads with mixed traffic and no separated bicycle infrastructure (1), bicycle lanes within the road way that are separated by a solid or dotted line (2), bicycle lanes physically separated from the roadway by a curb, parked cars or a green strip (3), and bicycle infrastructure that does not follow the roadway, such as paths in green areas (4). Within UR:BAN observational data is collected at urban intersections that are comprised of roads with different types of bicycle infrastructure and are signal controlled.

The three test intersections were selected based on the type of bicycle infrastructure available for bicycle traffic (types 1-3), the traffic flow and the possibility of installing a camera as high and as directly above the intersection as possible. Signal controlled intersections were selected for analysis for three reasons. First, many of the applications currently being developed in UR:BAN will be implemented at large, busy intersections, which tend to be signal controlled. Second, the design and evaluation of large, signalized intersections is more often carried out using microscopic traffic simulation tools, which will be enriched by the introduction of valid bicycle models. Finally, large intersections are particularly dangerous locations for bicyclists. Once potential intersections were identified, permission was gained from the public or private owner of the buildings used for camera mounting. The characteristics of each of the test intersections are given in Table 1.

Video data was collected using a GoPro Hero3 high definition video camera. This camera is inexpensive, small and can be discretely installed at intersections in such a way that most road users will not see that they are being filmed and therefore act differently than they normally would. In addition, the camera is enclosed in a waterproof case, which makes it possible to leave the camera outdoors. Video data was collected at 25 fps, which is the value recommended by the developers of “Traffic Intelligence”[4], with a resolution of 1920x1080. The use of a wide angle setting made it possible to collect information from the entire intersection as well as a section of the approaches from each direction. The camera was mounted using a theft-proof box in cases where it was accessible by the public and using a tripod in privately accessible areas. The personal privacy of the road users was respected as it was not possible to discern individual people or license plates in the video data.

Table 1: Characteristics of the Intersections

<table>
<tr><th colspan="2">Intersection 1: Arcisstraße and Theresienstraße</th></tr>
<tr><td>Description
</td><td><u>Theresienstraße</u>
• Three lanes of one way mixed traffic (type 1)
• Relatively high volume in the peak hours with comparatively low bicycle traffic
<u>Arcisstraße</u>
• One motor vehicle lane and one bicycle lane within the roadway in both directions (type 2)
• Traffic volume is low-medium with moderate bicycle traffic
Many pedestrians use this intersection as it is borders a university building.</td></tr>
<tr><td>Approximate Camera Height</td><td>14 m</td></tr>
<tr><th colspan="2">Intersection 2: Luisenstraße and Karlstraße</th></tr>
<tr><td>Description
</td><td><u>Luisenstraße</u>
• One lane for motorized traffic and a bicycle lane separated from the roadway by a curb in both directions (type 3)
• Motorized traffic volume is moderate and the bicycle traffic is relatively high because this is one of the main routes with a bicycle lane from the main train station to the university
• Several bus services run along this road
<u>Karlstraße</u>
• One lane with mixed traffic in both directions (type 1)
• The volume of motorized vehicles and bicycles is low
Many pedestrians use this intersection as it edges a college.</td></tr>
<tr><td>Approximate Camera Height</td><td>21 m</td></tr>
<tr><th colspan="2">Intersection 3: Seidlstraße and Marsstraße</th></tr>
<tr><td>Description
</td><td><u>Seidlstraße</u>
• Two lanes of through motorized traffic with a dedicated left turning lane in both directions and a dedicated right turning lane on the south arm of the intersection
• Bicycle lanes within the roadway in both directions (type 2)
• Heavy motorized vehicle and bicycle traffic as this is one of the main connecting roads under the train tracks from south to north Munich
<u>Marsstraße</u>
• Two lanes of through motorized traffic with a dedicated left turning lane on the east arm of the intersection and a dedicated right turning lane on the west arm of the intersection
• Bicycle lanes separated from the roadway by a curb in both directions (type 3)
• Moderate to heavy motorized and bicycle traffic in both directions (particularly turning south on Seidlstraße)
Many pedestrians use this intersection as it edges a shopping mall.</td></tr>
<tr><td>Approximate Camera Height</td><td>35 m</td></tr>
</table>

Calibration of Tracking and Grouping Parameters within "Traffic Intelligence"

"Traffic Intelligence" analyses video data in two steps. First, all the features in the video that move more than a defined minimum distance are tracked from one video frame to the next. In a second step, these tracked features are grouped together into hypothetical road users. The parameters associated with the first step in the algorithm control the number and quality of the features that are tracked. The parameters associated with the second step include the minimum number of and distance between features as well as the similarity between the trajectories of the features that are included into one road user hypothesis. There are a total of 22 parameters used by "Traffic Intelligence", 16 tracking parameters and six grouping parameters[7]. These parameters were adjusted logically and systematically until acceptable road user trajectories were derived. The findings of this calibration are described separately in the following two sections.

Feature Tracking Parameters

The tracking parameters included in "Traffic Intelligence" control the quality and dynamic characteristics of the features that are tracked. There are a total of 16 parameters used to control this process. In order to calibrate the software, all of the parameters were adjusted in a first step to discern whether or not they had a large effect on the outcome of the analysis, and if so, what effect was observed. Using this subjective method, six parameters were identified that had a considerable effect on the feature tracking results (shown in Table 2). The two most influential parameters, the maximum number of features tracked at one time and the minimum quality of the feature, were increased (1200) and decreased (0.005) respectively in order to ensure that a sufficient number of features were tracked. This seemed to have the most drastic impact on the results of the feature tracking software. In order to optimize the remaining parameters for the three test intersections, each was adjusted systematically until the best possible value for that parameter was determined. This was done by setting the parameter to an extreme low (or high) value and then increasing (or decreasing) the value until a critical point was reached when changing the value further decreased the quality of the results. The quality of the feature tracking results was evaluated using two criteria:

1. Maximize the portion of vehicles, bicycles and pedestrians that are tracked with a minimum of three features.
2. Minimize the time difference between the moment the road user begins moving and feature tracking.

It was found that an identical parameter set, given in Table 2, provided acceptable results for all three of the test intersections. One of the most important differences in tracking the features of bicycles in comparison with motorized vehicles concerns the acceleration of the features. While the speed difference between bicycles, pedestrians and motorized vehicles is not as pronounced at urban intersections as on road segments, bicycles and pedestrians accelerate slowly in comparison with motorized vehicles. In order to track the features of bicycles from the point that they start moving, the parameter minimum displacement to keep features, was set considerably lower than required for motorized traffic.

Table 2. Calibrated Tracking Parameters used for the Research Intersections

Parameter	Value
Maximum number of features tracked	1200
Minimum feature quality	0.005
Size of the search window at each pyramid level	8
Displacements to test minimum feature motion	2
Minimum displacement to keep features	0.005
Maximum feature acceleration	4

The feature tracking capabilities of "Traffic intelligence" are very good, even before calibration for the individual intersections. By slightly adjusting the identified influential parameters, it was generally possible to track at least one or two features of most of the road users. Still, features of the odd pedestrian or bicycle were not detected by the software.

Feature Grouping Parameters

The output of the calibrated feature tracking stage of the analysis is a large number of trajectories that follow moving features from one video frame to the next. In most cases, at least one, and often many, features of all road users are tracked while the object crosses the intersection. In a second stage of analysis, the tracked features are grouped into object hypotheses. This is done by considering the duration a feature is tracked, the minimum number of features required to make an object hypothesis, the similarity of the trajectory directions, and the spacing between the features throughout the time in which they are tracked. Although it seems relatively straightforward to group features into objects, the high volume of traffic at these intersections, the high physical and dynamic heterogeneity of the road users, as well as the similar trajectories observed when two road users travel next to each other in the same direction, made the calibration of this analysis stage much more challenging.

The physical differences between the types of road users at the analyzed intersections made the use of a common parameter set for the grouping of features nearly impossible. Not only are cars and heavy duty vehicles larger, but they also tend to have smoother lines, and therefore fewer edges that can be automatically tracked with the software. During the analyses it was observed that the tracked features of cars were often clustered around the front of the vehicle, particularly around the side view mirror, and around the end of the car. It is very difficult to use an algorithm to distinguish between two groups of features tracked on one car or those tracked on two bicycles riding single file. Bicycles and pedestrians on the other hand are not as smooth and have many corners and edges that can be detected and tracked by the feature tracking algorithms. The trajectories of these features however are less similar and may not travel in parallel, like those of motorized vehicles. Often the features of a foot of a bicyclist are tracked while they move in a circular pattern. The resulting trajectory is quite different than the features tracked on the head or body of the bicyclist. A similar situation occurs for the swaying arms of pedestrians.

In addition to the physical difference between road user types, there are also considerable differences in the dynamic characteristics. As the observations are made at intersections, the speed of the road user types is not significantly different. However, pedestrians and bicyclists have a maximum speed that they typically do not surpass. The acceleration rate of pedestrians and bicyclists is much lower than that of motorized vehicles.

Due to these differences, it was not possible to create a parameter set that allowed for the accurate grouping of features into road user hypotheses for all of the different road user types. For this reason, an approach was developed for separating the tracked features into two databases depending on the location of the moving features before the grouping of the road users was carried out. This was done by identifying regions of the video frame that are only touched by certain road user groups. Although bicycles and pedestrians also travel on the road surface, both in crosswalks, on bicycle lanes and in regions outside these designated areas, it is quite rare that a motorized vehicle drives on the sidewalk. Similarly, although a car may touch a bicycle lane that is within the street bounty, it is very unlikely that it will spend the duration of its crossing within this area. Based on these two assumptions, the study areas were divided into subareas by defining polygons. The first group of polygons, touch-at-any-instant polygons, encompasses sidewalks and areas that cannot be entered by motor vehicles. If a feature touches one of these polygons at any point during its existence, that feature is transferred to the "bicycle/pedestrian" feature database. The second group of polygons, touch-in-all-instances polygons, encompasses bicycle lanes. If a feature exists within the polygon for the entire duration that it is tracked, it is assumed to be a bicycle and that feature is transferred to the bicycle/pedestrian feature database. Originally it was intended to include touch-in-all-instances polygons for crosswalks, but it was found that many right turning vehicles start almost directly at the edge of the crosswalk, drive through the crosswalk, and then stop to wait for bicycles or pedestrians in the parallel direction. The result was that many cars were falsely identified as pedestrians crossing the sidewalk perpendicularly. However, since almost all pedestrians touch the sidewalk at some point, the touch-in-all-instances polygons were found to be sufficient for capturing pedestrian features. An example of the polygons overlaid over a video frame for the intersection Marsstraße-Seidlstraße is shown in Figure 1. The other two intersections were divided similarly.

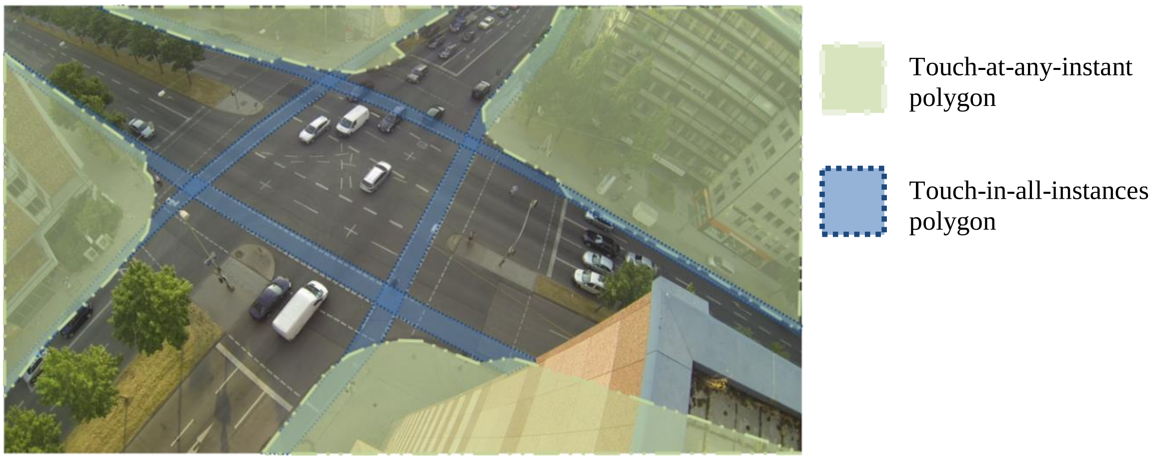


Figure 1: Example of Polygons overlaid over Video Frame (Marsstraße-Seidlstraße)

Once two separate feature databases were created, the features were grouped into road user hypotheses using parameter sets adjusted for the physical and dynamic characteristics of those groups. The parameters used to control the grouping of features are given in Table 3 and the calibrated values used to group features at the three intersections for the pedestrian/bicycle database and motorized vehicle database are given.

The values for the three intersections were arrived at in a similar manner as those for the feature parameter analysis step. Each of the six grouping parameters in “Traffic Intelligence” were adjusted independently, starting first at an extreme value (high or low) and then increasing or decreasing that value until the optimal results were achieved on a one minute test video clip. The quality of the results was determined qualitatively by assessing the number of over segmented or over grouped road users. Over grouping refers to the combination of more than one road user into one object, while over segmentation occurs when the features from one road user are grouped as more than one road user. However, if the parameters are adjusted to reduce the occurrence of over grouping, the number of over segmented objects will tend to increase. Inversely, if over segmentation is minimized, the occurrence of over grouping will increase. This is particularly true when objects with different physical and dynamic characteristics are analyzed together. Although the separation of the database into two databases with objects of similar size made it possible to correctly group features and minimize the occurrence of over grouping and over segmenting, such errors could not be completely avoided. Because the manual correction of over segmentation is much easier than that of over grouping, parameters were set to favor over segmentation.

Table 3: Calibrated Grouping Parameters for Bicycles and Pedestrians and Motorized Vehicles

Parameter	**Arcisstraße**	**Luisenstraße**	**Seidlstraße**
	(Bicycles and Pedestrians/Motorized Traffic)		
Minimum number of frames to consider a feature for grouping	40/45	40/40	40/30
Connection distance (distance at first instant) (m)	0.8/1.9	0.8/2.1	0.8/2.0
Segmentation distance (different between maximum and minimum distance) [m]	0.4/0.7	0.5/0.6	0.5/0.5
Maximum distance [m]	0.8/1.9	0.8/2.1	0.8/2.0
Minimum cosine of the angle between the velocity vectors for grouping	0.65/0.70	0.6/0.65	0.6/0.7
Minimum average number of features per frame to create a vehicle hypothesis	1/3	1/1	1/1

2.2 Extension 1: Merging of Disconnected Trajectories

As “Traffic Intelligence” tracks features that are moving between video frames and then creates a road user hypothesis from these tracked features, a road user who stops moving is no longer tracked. When the road user begins moving again, it is assigned a new object identification number and the trajectory data is stored separately for both objects, before and after the stop. Additionally, the trajectory of a road user is often interrupted if the road user passes behind or under and obstructive object, such as an overhead cable hanging between the video camera and the intersection, a sign post or the pole of a traffic light. When this occurs, the road user is tracked as two objects simultaneously as it passes behind or underneath the obstacle, and then the first identification number of the object is discontinued.

The original purpose of “Traffic Intelligence” was to analyze traffic safety, which does not require complete trajectories of road users as the cross an area. However, in order to analyze tactical behavior, it is necessary to be able to track a road user throughout its entire path across an intersection. To accomplish this, the data output from “Traffic Intelligence” was processed again using an algorithm that searches for objects that end in the middle of the video frame. When such instances are found, a search is carried out for objects that appear at a later point in the video sequence within a predefined radius of where the first object stopped being tracked. The trajectories of these two road users are then merged. In order to identify cases where an object is tracked as two objects simultaneously as it passes behind an obstacle, the algorithm also searches for objects that appear at the same time or before the first object disappears.

The value used for the search radius must be adjusted in order to maximize the number of correct connections made while minimizing false connections. If the radius is too large, trajectories of different road users could mistakenly be connected. This can happen at stop lines where many objects start and stop at nearly the same point. However, if the search radius is set too small, trajectories belonging to the same road user may not be merged. The tradeoff between setting a large radius and merging trajectories that do not belong to the same road user and selecting a small search radius and not merging trajectories that belong together is similar to that of setting the correct grouping parameters to minimize over grouping and over segmentation. In this case, the radius was set relatively small (30 cm), decreasing the number of falsely connected road users but increasing the required manual processing effort afterward.

The merging algorithm did not attempt to correct instances of over segmentation, when one road user is tracked as two or more road users, as this was already done by adjusting the grouping parameters in “Traffic Intelligence”. However, as the parameters were set to favor over segmentation rather than over grouping, a number of road user trajectories must be merged manually after the connecting algorithm is implemented.

2.3 Extension 2: Classification of Road Users

Once the segmented trajectories of all the road users are connected, the road user is categorized based on its trajectory as it crosses the intersection. Logical assumptions are made based on the geometry of the intersections and the traffic rules. This was done in the same way as for the separation of features into two databases. It is assumed that if a road user touches a sidewalk at any point during its existence, it is either a bicycle or a pedestrian. In addition, if there is a median in the roadway, small bicycle detecting polygons are placed in front of the median, where motor vehicles are unlikely to cross. If the road user does not touch any of these polygons at any point while it crosses the intersection, it is categorized as a motor vehicle. Differentiation of motorized vehicles types, including cars, motorcycles and trucks, was not attempted in this work as the focus is on bicycle traffic, but would be useful to do in the future.

If a road user does touch one of the polygons, it is classified as a bicycle or pedestrian. A simple approach for differentiating between bicycles and pedestrians based on the maximum speed of the object was implemented. A maximum speed for pedestrians of 9 km/h and a maximum speed of bicycles of 30 km/h were adopted. These speed limits were selected based on the observed speeds of the bicycles and pedestrians in the video data from the three crossings. If an object in the pedestrian/bicycle database had a maximum speed of over 30 km/h, it was assumed to be a falsely detected motor vehicle and was deleted from the database.

2.4 Analysis of the Resulting Trajectory Data

The resulting trajectory data contains the position and speed of each pedestrian, bicycle and motor vehicle in each video frame (25 fps) for its entire pathway across the intersection. This dataset will be used to analyze the movement and interaction of bicycles at urban intersections, both at the operational and tactical level, to provide necessary input for the calibration and validation of existing bicycle models and to create new models where necessary. At the operational level or control level, which includes the automatic action patterns used by a road user to maneuver through the road environments[3], a number of models currently exist and can be used to depict bicycle traffic. Examples of movement and interaction at this level include adaptation of speed based on the road alignment or the movement of other road users. However the majority of these models are not yet calibrated or validated. The required information to validate and calibrate exiting models includes the speed and acceleration profiles of bicycles using different infrastructure types and performing various maneuvers (e.g. riding straight, turning left, and turning right) and the lateral and longitudinal spacing as well as the density (bicycles/m^2) maintained by cyclists while riding and queuing will be derived from the trajectory database produced in this project. At the tactical level or maneuvering level, which includes controlled action patterns[3] such as the selection between available infrastructure (bicycle lane, sidewalk or road), far fewer models are available. The trajectory data will be combined with situational data, such as the traffic volume, the position and speed of nearby road users and the phase of the signal control in order to analyze the tactical selection of 'route choice' across the intersection.

The analysis of the tactical selection of a route across the intersection requires input regarding the maneuver and infrastructure used by the bicyclist. Due to the large size of the dataset, an automated method for classifying the different driving and riding maneuvers carried out by cars and bicyclists respectively must be implemented. The eleven driving maneuvers identified by Dambier[8] were used as a starting point for the classification carried out in this research. Each of these maneuvers is mathematically defined for car drivers. These equations have been adapted as far as possible. A number, however, had to be redefined due to the lack of mechanical information, including for example braking pressure and angle of the steering wheel. This methodology is currently being implemented.

3. RESULTS

The output from the calibrated "Traffic Intelligence" software is analyzed to determine the percentage of road users that are correctly identified and tracked in three minute sample video segment from each of the three test intersections. The accuracy of the road user classification algorithm is evaluated for cars, bicycles and pedestrians at the three intersections. The trajectory merging algorithm is tested using the video data sample from the intersection Luisenstraße-Karlstraße. The effort required to manually correct the trajectory data after the two extension algorithms have been run is estimated.

3.1 Assessment of the Tracking and Classification of Road Users

The output from the calibrated "Traffic Intelligence" was assessed based on the percentage of each type of road user that was tracked or missed by the software. The number of instances that road users that were over grouped, meaning that two or more road users are grouped together and classified as one road user, were also counted. As it is not possible using this approach to segment over grouped road users in a secondary step, these inaccurate trajectories are deleted from the database. The over grouped road users are then counted as missing in the analysis. The number of over grouped pedestrians was not analyzed because this often occurs when two pedestrians walk side by side, which happens very frequently at urban intersections. This over grouping was deemed to be unproblematic, however, because the operational and tactical behavior of these pedestrian pairs is identical. Cases of over segmentation were also not counted because these also occurred quite frequently and were corrected manually using a database editing tool. The time required to manually edit these files has been noted.

To assess the algorithm used to classify each road user as a car, bicycle or pedestrian, a three minute sample of video data from each of the three intersections was analyzed. The actual number of each type of road users was manually counted in the video segment. The number and percentage (of road users that were tracked by "Traffic Intelligence") of each type of road user at each intersection that were correctly classified was recorded. The accuracy of the object tracking by "Traffic Intelligence" and the classification methodology are shown in Table 4.

Table 4: Accuracy of the Tracking by “Traffic Intelligence” and the Classification Methodology

Road User/Intersection	Total Number	Total Tracked (%)	Over Grouping # of instances (%)	Correctly Classified of those tracked (%)
Motorized Vehicles:				
Arcisstraße	57	41 (72%)	2 (7.0%)	39 (95%)
Luisenstraße	36	33 (92%)	1 (5.5%)	30 (91%)
Seidlstraße	111	95 (86%)	2 (3.6%)	92 (97%)
Bicycles:				
Arcisstraße	25	18 (72%)	0 (0.0%)	5 (28%)
Luisenstraße	25	19 (76%)	1 (8.0%)	11 (58%)
Seidlstraße	30	24 (80%)	0 (0.0%)	23 (96%)
Pedestrians:				
Arcisstraße	14	13 (93%)	-	8 (62%)
Luisenstraße	23	18 (78%)	-	14 (78%)
Seidlstraße	17	13 (76%)	-	8 (62%)

The accuracy of the methodology used to track and classify cars, bicycles and pedestrians varied widely between the three research intersections. The percentage of tracked road users ranges from 72%-93%, while the percentage of correctly classified road users ranges from 28%-97%. The most accurate results were obtained for cars and bicycles at the intersection Seidlstraße-Marsstraße. A number of factors contributed to the better results obtained at this intersection. First, the camera was installed at a very high vantage point with a very high angle to the intersection, which enables better tracking of moving features[4]. Second, the existence of bicycle lanes in all directions, as well as the high use of this infrastructure by bicyclists made the hypothesis of road user type based on the position of the moving features quite accurate. Finally, the relatively high volume of vehicular traffic was found to restrict bicycle movement and behavior to predictable patterns (use of bicycle lanes, crossing at intended times and locations), which again made the classification of road user type more accurate.

The results obtained at the intersection Luisenstraße-Karlstraße were quite good for motorized vehicles and pedestrians, but the misclassification of bicycles as cars lead low accuracy in bicycle classification. This is largely due to the fact that bicycles travel on Karlstraße with the motorized vehicles (no separated bicycle infrastructure) and were therefore not transferred to the bicycle database. This factor likely had the additional result of increasing the percentage of bicycles that were missed because the grouping parameters used to analyze the car database were not optimized for bicycle traffic. This also occurred at the intersection Arcisstraße-Theresienstraße, where the bicycles ride with motorized vehicle traffic on Theresienstraße.

The percentage of missed bicycles and pedestrians at the intersection Arcisstraße-Theresienstraße (28%) is considerably higher than at the other two intersections. This could be due to the low vantage point (14 m) and the low angle between the camera and the intersection. Another problem at this intersection is the presence of a cable that is used to hang a street light over the middle of the intersection that hang directly between the camera and the intersection. This cable caused considerable problems for the grouping and classification of road users. Additionally, during data collection there was a construction site on Arcisstraße that blocked the use of the bicycle lane in the north approach. As a result, the majority of the bicyclists used the road with the motorized traffic and were therefore categorized as cars.

The percentage of missed bicycles is slightly higher at all three intersections than the percentage of missed cars and pedestrians. This is likely because bicycles travelling faster than 30 km/h were erased from the bicycle/pedestrian database in order to reduce the number of cars falsely detected as bicycles. The percentage of missed bicycles could be slightly reduced if the objects with a speed over 30 km/h were not removed from the pedestrian/bicycle database. This however would also result in an increased percentage of cars that are falsely identified as bicycles or pedestrians.

The manual correction of the road user type of misclassified objects using a database editing tool created for this project took between 1-2 minutes per minute of video data. As stated before, it was not attempted to manually trace the trajectories of road users that were missed by “Traffic Intelligence”.

3.2 Assessment of Combination of Trajectories

The automated method developed to merge segmented trajectories belonging to the one road user proved to be useful in reducing the amount of manual processing required to create complete trajectories of the road users as they cross the intersection. The merging method was evaluated using the video sample from the intersection Luisenstraße-Karlstraße. Of the required 130 merges, 114 (88%) were completed using the logical algorithm. Five false merges, where trajectories belonging to different road users were merged together, were also carried out. It was noted that the algorithm was more successful in merging trajectories in situations where the road user stopped and started again. Merges were often missed when an object passed behind or under and obstructing obstacle.

The manual merging of trajectories that were missed by the merging algorithm took between 1-3 minutes per minute of video data depending on the traffic volume at the intersection. The manual merging of over segmented road users was found to be more time consuming because the object identification numbers often overlap each other, making it difficult to determine which trajectories must be merged. For the sample video at the intersection of Luisenstraße-Karlstraße, an additional 72 manual merges were necessary to combine the trajectories of over segmented road users, which took about 8 minutes to identify and enter into the database editing program. In general, the combining of over segmented road users was found to take between 2-4 minutes per minute of video data.

4. DISCUSSION

The methodologies presented in this paper offer a good starting point for the classification of road users at urban intersections and the connection of segmented trajectories based on the open source software “Traffic Intelligence”. Depending on the physical characteristics of the research intersection, the volume of bicycle, pedestrian and motor vehicle traffic, the height and angle to the intersection of the video camera installation and the presence of any obstructing objects such as overhead wires, up to 97% of the motorized vehicles, 96% of bicycles and 78% of pedestrians are correctly classified using the developed algorithm. The merging algorithm was found to identify and merge 88% of the necessary merges. However, the resulting dataset still contains over segmented, over grouped, misclassified, and missed road users. The over segmentation and misclassification can be corrected after the automated analysis by watching the videos, manually noting which road users must be combined or reclassified, and changing this in the database using a tool developed for that purpose. It has been found that the time required to manually correct the resulting data takes approximately 4-9 minutes for each minute of video data, depending on the traffic volume at the intersection. The correction of missed or over grouped road users is, however, much more difficult and was not done in this work. While the percentage of missed motor vehicles is relatively low, a considerable number of bicycles and pedestrians are not tracked by the software, despite the effort made to calibrate “Traffic Intelligence”. The resulting data base is then lacking 7-28% of the pedestrian and bicycle traffic. Future work will aim to improve these numbers. Nevertheless, care must be taken when using the data base to make conclusions about bicycle behavior, as the missing road users could have had an influence on the behavior of the analyzed bicycle.

The position of the video camera and the presence of overhead cables and other obstructive objects were found to have a very large impact on the percentage of road users tracked. The percentages of cars and bicycles tracked at the intersection Arcisstraße-Theresienstraße are considerably lower than at the other two intersections. This could be due to the comparable low vantage point (14 m) and the correspondingly low angle between the video camera and intersection. Furthermore, the proximity of the overhead cables to the video camera caused problems with the segmentation of road user trajectories. The selection of intersections with a vantage point over 20 m high, with as few obstructing objects is therefore recommended for data collection.

Further possibilities to extend the presented methodologies are currently being implemented. In order to increase the reliability of the differentiation between pedestrians and bicycles, the acceleration will be considered in addition to the maximum speed. An analysis of the number of features and feature density associated with each type of road user will also be carried out to determine if there is any correlation that could be used in combination with the maximum speed and acceleration. Finally, the possibility of using a Fuzzy Logic based approach to determine the type of road user will be developed and tested.

A problem with the use of polygons to identify regions of the intersection occurs at the borders of each polygon. If a road user travels very near to the boundary of the polygon, its features tend to be divided into both the car and bicycle/pedestrian database. This leads to the creation of two road user hypotheses, a car and a bicycle or pedestrian, when the two databases are recombined. In order to reduce this effect, a narrow void area will be introduced at the borders of the polygon where all features are deleted. Thus the few features of a car that pass over the boundary into the bicycle/pedestrian area will be deleted and not formed into a vehicle hypothesis, and vice versa.

This approach is based on the assumption that certain road users tend to travel on certain parts of the road infrastructure. However, this is not always the case. Shared space situations, where cars, bicycles and pedestrians move freely in the given space at walking speed, would be very difficult, if not impossible to analyze using this approach. Even at intersections with separated infrastructure for cars, bicycles and pedestrians, the accuracy of this methodology depends on the predictability of the road user behavior. At intersections with lower vehicular traffic, pedestrians and bicycles tend to cross at unexpected locations, making the correct classification impossible with this approach. In order to be able to analyze intersection with low vehicular traffic, other possibilities for automated video analysis must be developed. These may include the use of heat sensor cameras and the recognition of shapes, such as bicycle wheels.

5. CONCLUSION

The presented approach for extracting road user trajectories from video data using “Traffic Intelligence”, classifying the tracked road users as pedestrians, bicycles or motor vehicles and combining segmented trajectories proved to be effective in reducing the required analysis time while still producing sufficiently accurate results. The dual calibration of the software to consider the unique physical and dynamic characteristics of the different types of road user groups made it possible to minimize over grouping of bicycles and pedestrians, or conversely, the over segmentation of motor vehicles. The manual correction of over segmented road users is possible with relatively little time and effort (2-4 minutes per minute of video data). The use of polygons to specify regions that are only touched by certain types of road users in combination with the maximum speed of the road user to create hypotheses regarding the type of road user proved to be useful, although not perfect. The reassignment of falsely classified road users was found to be minimal (1-2 minutes for one minute of video data.)The automated merging of disconnected trajectories was found to connect 88% of the disconnected trajectories. The effort required to manually merge the remaining 12% was found to be minimal (1-3 minutes per minute of video.)

The resulting database consists of trajectory data from classified road users as they cross the complete intersection. This will then be used to derive information concerning the operational and tactical behavior of bicyclists at urban intersections. Besides the planned use of this data for the calibration, validation, extension and creation of bicycle models for microscopic traffic simulation tools, a great deal of additional research will be possible with the data.

6. ACKNOWLEDGEMENTS

This research is supported by the Federal Ministry of Economics and Technology on the basis of a decision by the German Bundestag. Research was carried out within the framework of the project UR:BAN (Urban Space: User oriented assistance systems and network management).

Thank you to Nicolas Saunier and his team for their development and open source sharing of “Traffic Intelligence” and for the support in using and extending the software.

Original Citation:
Heather Twaddle ; Tobias Schendzielorz ; Oliver Fakler and Sasan Amini, "Use of automated video analysis for the evaluation of bicycle movement and interaction", Proc. SPIE 9026, Video Surveillance and Transportation Imaging Applications 2014, 90260U (March 5, 2014).

http://dx.doi.org/10.1117/12.2042487

7. REFERENCES

[1] Hütter, A., "Verkehr auf einen Blick", Statisches Bundesamt, Wiesbaden, (2013).

[2] Twaddle, H., Schendzielorz T, Fakler O., "Bicycles in urban areas: Review of existing methods for modelling behavior," Proc. Transportation Research Board 93rd Annual Meeting, (2013).

[3] Michon, J. A., [A Critical Viel of Driver Behavior Models: What do we know, what should we do?], Plenum Press, New York, 485-520 (1985).

[4] Jackson, S., Miranda-Moreno, L., St-Aubin, P., Saunier, N., "A flexible, mobile video camera system and open source video analysis software for road safety and behavioural analysis," Transportation Research Board Annual Meeting Compendium of Papers, 13-3229 (2013).

[5] Bördlein, E.,"The Munich Bicycle Development Concept,"Planungsreferat (City Development Department), Munich, (2000).

[6] Monheim, H., Muschwitz, C., Philippi, M.,"Evaluationsbericht der Fahrradmarketingkampagne Radlhauptstadt München,"Wuppertal Institut,(2011).

[7] Saunier, N., “Traffic Intelligence,” 12 November 2013, <https://bitbucket.org/Nicolas/trafficintelligence-/wiki/Home> (20 December 2013).

[8] Dambier, M., [Adaptive Information Flow Control – Recognition and Prediction of Factors contributing to Driver’s Stress], Universität der Bundeswehr München, Munich, (2010).

[9] Zangenehpour, S., Miranda-Moreno, L., Saunier, N., "Impact of bicycle boxes on safety of cyclists: a case study in Montreal", Transportation Research Board Annual Meeting Compendium of Papers, 13-2909 (2013)

[10] Li, S., Sayed, T., Zaki, M.H., Mori, G., Stefanus, F., Khanloo, B., Saunier, N., "Automatic collection of pedestrian data using computer vision techniques,"Transportation Research Record: Journal of the Transportation Research Board 2299, 121-127 (2012).