

MEASURING TRAFFIC FLOW AND LANE CHANGING FROM SEMI-AUTOMATIC VIDEO PROCESSING

Marcel Sala Sanmartí

Research fellow, CENIT - UPC-BarcelonaTech, Spain

Francesc Soriguera Martí

Associate Professor, CENIT - UPC-BarcelonaTech, Spain

ABSTRACT

Comprehensive databases are needed in order to extend our knowledge on the behavior of vehicular traffic. Nevertheless data coming from common traffic detectors is incomplete. Detectors only provide vehicle count, detector occupancy and speed at discrete locations. To enrich these databases additional measurements from other data sources, like video recordings, are used. Extracting data from videos by actually watching the entire length of the recordings and manually counting is extremely time-consuming. The alternative is to set up an automatic video detection system. This is also costly in terms of money and time, and generally does not pay off for sporadic usage on a pilot test.

An adaptation of the semi-automatic video processing methodology proposed by Patire (2010) is presented here. It makes possible to count flow and lane changes 90% faster than actually counting them by looking at the video. The method consists in selecting some specific lined pixels in the video, and converting them into a set of space – time images. The manual time is only spent in counting from these images. The method is adaptive, in the sense that the counting is always done at the maximum speed, not constrained by the video playback speed. This allows going faster when there are a few counts and slower when a lot of counts happen.

This methodology has been used for measuring off-ramp flows and lane changing at several locations in the B-23 freeway (Soriguera & Sala, 2014). Results show that, as long as the video recordings fulfill some minimum requirements in framing and quality, the method is easy to use, fast and reliable. This method is intended for research purposes, when some hours of video recording have to be analyzed, not for long term use in a Traffic Management Center.

1 INTRODUCTION

The information given by different freeway sensors, such as loop detectors or License Plate Recognition (LPR) devices, is very valuable, but still incomplete in order to address some research questions. The detailed measurement of vehicles' trajectories would be the desired and most complete information. However, this is not generally available for all vehicles in a traffic stream, and researchers need to use ad-hoc video recordings to complete their databases with the required information.

In such situations, video processing, aimed to extract any type of data from the recordings, appears as an issue. Full automatization is not usually feasible. The limited amount of data to treat in a research pilot test does not justify the implementation of complex automatic image processing systems. Researchers usually face the task with labor, knowing that visually extracting data from video recordings is extremely tedious and time consuming.

Freeway lane changing activity is the required data to extract from videos in the present paper. Lane changes needed to be counted in six different freeway stretches for a total of 63 hours of video recordings. The development of a method for extracting this information in a fast, simple and reliable way is the aim of this paper.

The paper is structured as follows: Section 2 reviews different methods for video processing in order to extract traffic data; next, in Section 3 a detailed description of the new methodology for counting lane changes is presented. Recommendations to correctly use the method are also described; in Section 4 the method is applied to some video recordings taken at the B-23 freeway in Spain; Finally, in Sections 5 and 6 the potential of lane changing data is discussed, and some conclusions are outlined.

2 TRAFFIC VIDEO PROCESSING: A STATE OF THE ART

There are several methods to extract lane changes or other traffic information from a video recording. They range from fully automation, to visual inspection with handwritten notes. All of them present advantages and drawbacks. A selection of some methods of interest is presented next, starting from the most automatic ones.

2.1 NGSIM software

To the authors knowledge, the most automatic video detection tool used for traffic analysis is the NGSIM software (Federal Highway Administration, 2015; Federal Highway Administration, 2006). The NGSIM (Next Generation SIMulator) is a powerful software developed by Cambridge Systematics, Inc. for the Federal Highway Administration of the US government. The software is able to automatically detect every single vehicle in the study area, and follow its trajectory. It also detects the size of each vehicle. An example of usage along with the resulting dataset at I-80 location in California can be found at Punzo et al. 2011, and Lu & Skabardonis 2007.

However, there are several requirements in order to use the software. First, the software only works if using special (and expensive) cameras which need to be specifically prepared to that end. All cameras must be calibrated, and the video recordings stabilized, rectified and georeferenced. Because the software is not able to re-detect a particular vehicle in different cameras when it leaves the recording area, video recordings need to overlap. With no overlapping, its usage is limited to the coverage of one camera.

2.2 Multi-purpose video processing tools

OpenCV (Open Source Computer Vision - <http://opencv.org/>) is one of the most developed automatic image processing environments. This platform includes a set of general purpose video analysis tools for real time processing. Some of them have been used to count vehicles from video recordings (Uke & Thool, 2013). They work reasonably well, and developers claim a reliability around 90% to 95% (Barragán, 2015; Bhaskar, & Yong 2014). However, they are unable to automatically detect lane changes without human interaction.

Other advanced image processing techniques based on machine learning, focus on vehicle tracking using regular video recordings without any special camera requirement (Oliveira & Santos, 2008). While they are able to detect and count cars reasonably well in the proximity of the camera, problems appear in the vehicle tracking through the camera coverage, especially at the furthest part of the recording. The result is that the technique is

still not reliable enough to count lane changes, although progress is expected.

2.3 Semi-automatic video processing

Semi-automatic processing includes all the methods that cannot completely eliminate the human visual intervention, but that can limit it to very specific tasks, improving the efficiency with respect to the completely visual and manual processing. The methodology developed by Patire 2010 and Patire & Cassidy 2011, for a specific traffic analysis on the Tomei expressway accessing Tokyo in Japan, falls under this category.

Patire's implementation used 11 cameras spaced about 100 meters, without in between on/off-ramps. The method, first converts the video to still images, called epochs. An epoch is the image resulting from one pixel line of the video accumulated through time; this line is called "the scan line" (see Fig. 1). This implies that, among all the pixels of the video scene, only one line is used. From the epochs, and with a few manual clicks, it semi automatically recognizes the vehicles in successive cameras. By inferring between the discrete camera locations, it can be detected if a vehicle has changed lane. Also, the approximate vehicle trajectory can be obtained.

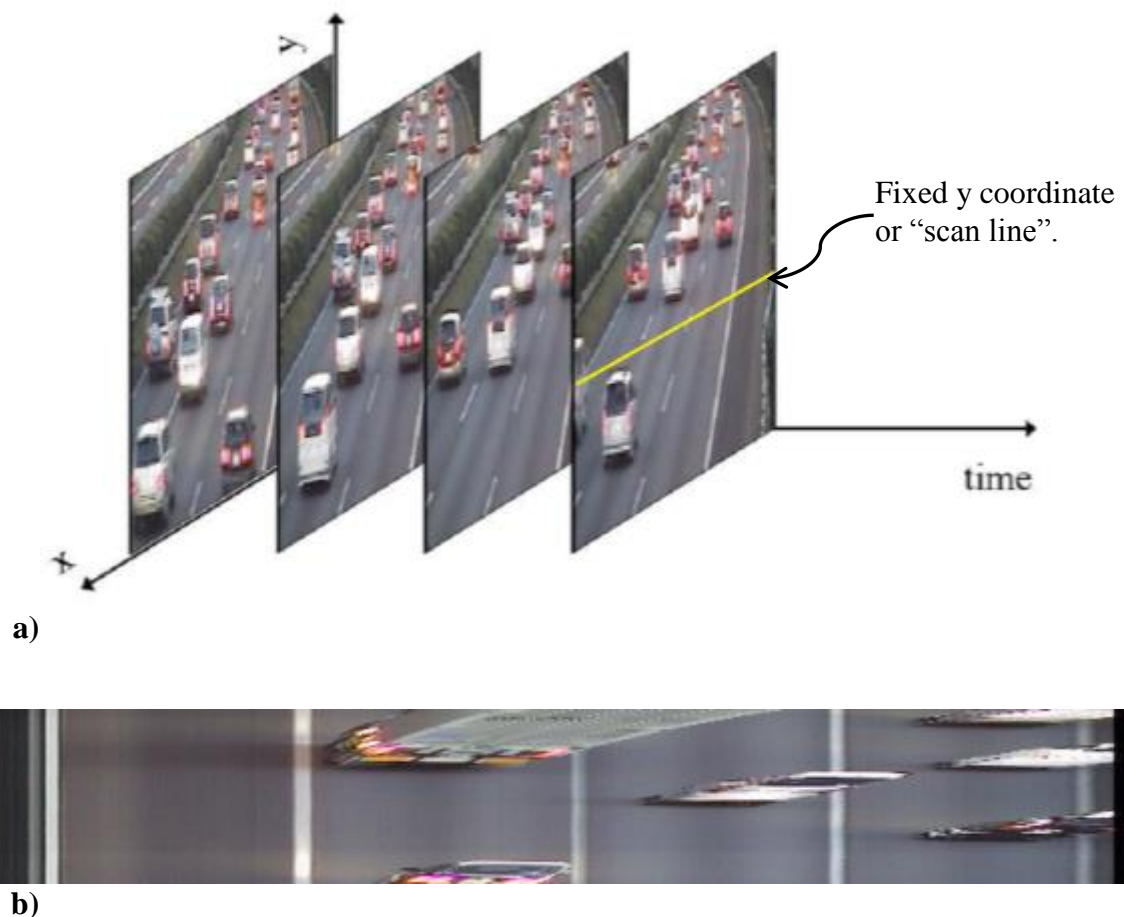


Fig. 1 - Transformation from video to epochs. a) Video frames over time. b) Epoch.
Adapted from (Patire 2010).

2.4 Enhanced visual video processing

Some options exist to enhance the completely manual procedure, for instance by automatically saving the counts when the user "clicks" or enabling to vary the video

playing speed. These options eliminate the need of note taking and can speed up the visual video processing, although the entire video length still needs to be played. Implementations for tactile devices (Campbell & Skabardonis, 2013; Campbell 2012) allow including (x,y) coordinate reference to the measured variable.

2.5 Completely visual video processing

This is the raw option considered as the baseline reference for comparison. It implies watching the entire video while taking notes of every lane change seen. Authors' experience confirms that lane changes can only be counted reliably by playing the video at a maximum of double speed. The entire video needs to be played for every pair of lanes.

3 NEW SEMI-AUTOMATIC VIDEO PROCESSING METHOD FOR MEASURING LANE CHANGING ACTIVITY

3.1 The new scan line

The methodology is based on the idea of the scan line described in (Patire, 2010). However, the scan line is no longer a straight horizontal line, but any line on the video scene. Therefore the pixel selection to construct the epoch is any set of coordinate points. An example of scan line used for measuring lane changing activity is shown in Fig. 2. Note that the scan line follows the lane division so that any vehicle crossing the line (i.e. a lane change) will appear as a spot in the epoch.

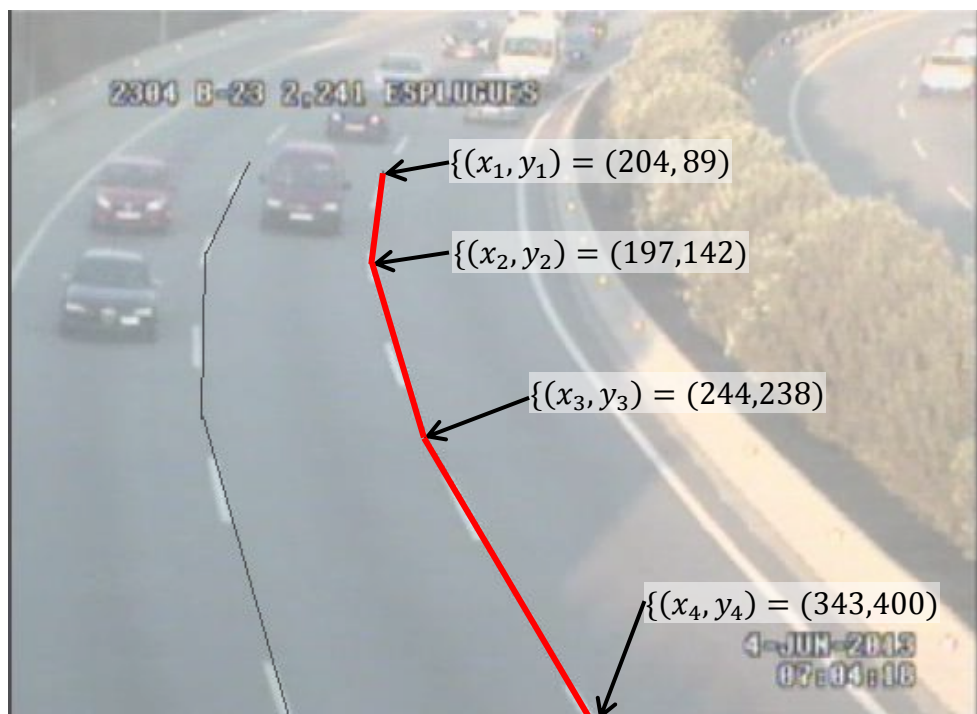


Fig. 2 - Sample scan line to compute lane changes between shoulder and central lanes. n=4 points.

Then, the inputs for the methodology simply are:

- The video file
- The scan line definition
- The desired epoch duration

Here, the scan line is assumed to be piecewise linear, defined by the pixels at the (n) breakpoints (see Equation 1). This assumption is not limiting in any sense and any type of line could be defined. Note that the units of the coordinate points (x, y) axis are pixels.

$$\{(x_i, y_i) \mid \forall i \in [0, n], x_i \in \mathbb{N}, y_i \in \mathbb{N}\} \quad (1)$$

Each segment of the scan line s_i is defined by a consecutive pair of points (2). Then, the coordinates of the ordinal set of pixels belonging to the segment can be computed according to Equations 3 and 4:

$$s_i \doteq \{(x_i, y_i); (x_{i+1}, y_{i+1}) \mid \forall i \in [1, n - 1]\} \quad (2)$$

$$y = a_i \cdot x + b_i \text{ s.t. } \begin{cases} y_i = a_i \cdot x_i + b_i \\ y_{i+1} = a_i \cdot x_{i+1} + b_i \end{cases} \quad (3)$$

$$\text{if: } |y_i - y_{i+1}| \geq |x_i - x_{i+1}|$$

$$(\vec{x}_i, \vec{y}_i) = \left(\left\| \frac{y - b_i}{a_i} \right\|, y \right) \begin{cases} y \in [y_i, y_{i+1}]; y \in \mathbb{N}, & i = 1 \\ y \in (y_i, y_{i+1}]; y \in \mathbb{N}, & i > 1 \end{cases} \quad (4)$$

else:

$$(\vec{x}_i, \vec{y}_i) = (x, \|a_i \cdot x + b_i\|) \begin{cases} x \in [x_i, x_{i+1}]; x \in \mathbb{N}, & i = 1 \\ x \in (x_i, x_{i+1}]; x \in \mathbb{N}, & i > 1 \end{cases}$$

Finally, the complete scan line is simply obtained by merging all the segments, as in Equation 5:

$$(\vec{x}, \vec{y}) = \bigcup_1^{n-1} (\vec{x}_i, \vec{y}_i) \quad (5)$$

And the number of pixels in the scan line (m) is computed by adding up the number of pixels per each segment (m_i) as in equations 6 and 7:

$$m_i = \max(|y_i - y_{i+1}|, |x_i - x_{i+1}|) + \begin{cases} 1, & i = 1 \\ 0, & i > 1 \end{cases} \quad (6)$$

$$m = \sum_1^{n-1} m_i \quad (7)$$

3.2 Constructing the epoch

Once the scan line is defined, it is possible to construct the epoch. The epoch size (8) depends on m (i.e. vertical dimension, corresponding to the number of pixels in the scan line), on the video frame rate (fps - frames per second) and on the epoch length in seconds t_e :

$$\text{Epoch size (horizontal} \times \text{vertical)} = w \times m \text{ [pixels]} \quad (8)$$

Where:

$$w = \text{round}(fps \cdot t_e) \quad (9)$$

The epoch duration, t_e , needs to be much longer than the typical lane-changing duration, minimizing the probability of one lane-changing maneuver being split over two or more epochs. For the present application (see Section 4), a duration of one minute was selected, this being long enough to satisfy the previous condition. In addition, this also corresponds to the traffic detectors aggregation period. The result of accumulating the scan line for each frame through time is the final epoch, as seen in Fig. 3.

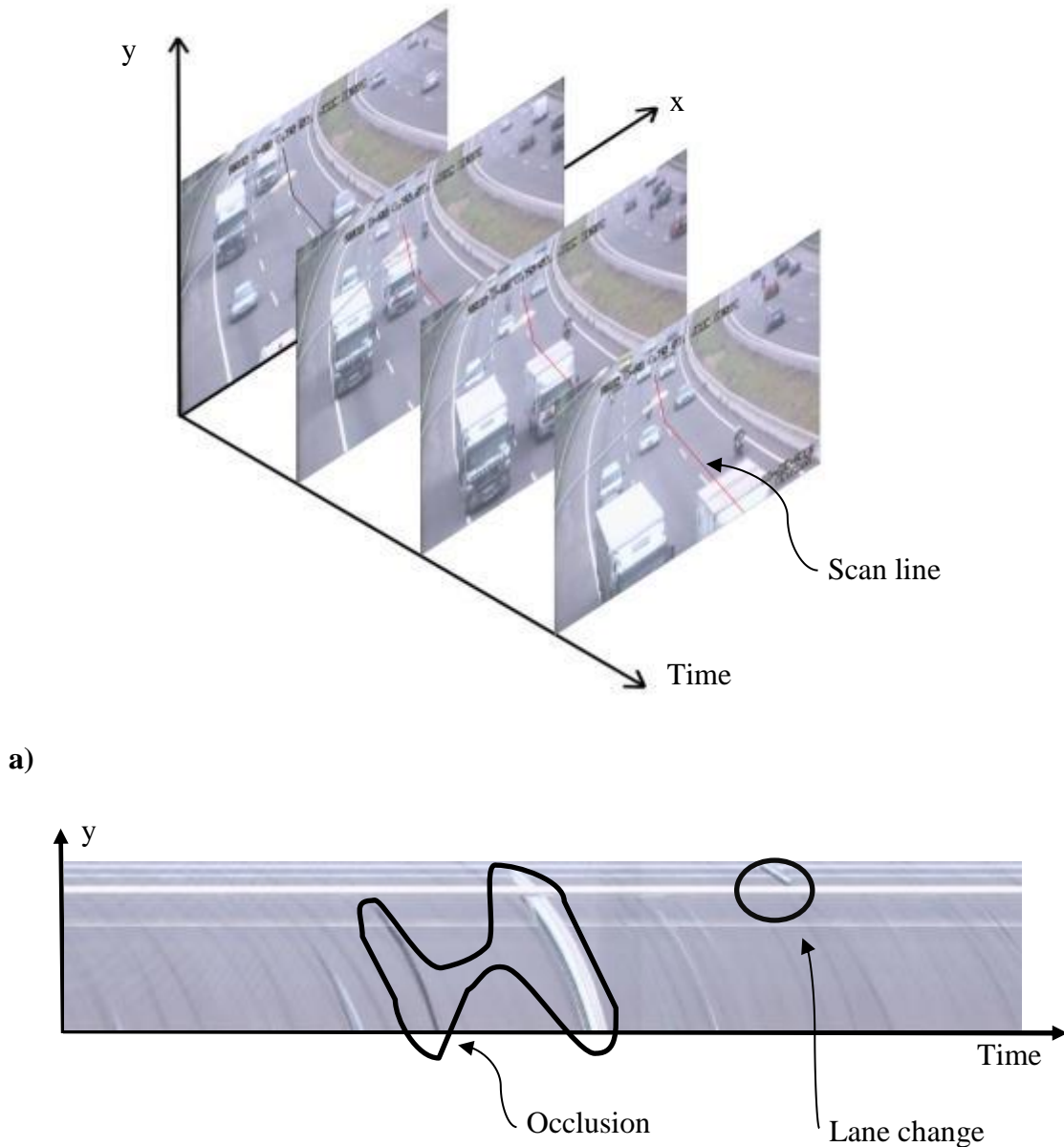


Fig. 3 - Epochs for semi-automatic lane change count. a) Epoch construction b) Sample epoch.

3.3 Lane change identification

As seen in Fig. 3, not only lane changes appear in the epochs, but also occlusions. This happens when high vehicles occlude the scan line. Occlusions can be easily differentiated from lane changes. They appear as big long shapes, covering a significant part of the scan

line. The shape and position of the occlusion varies depending on the camera framing and the scan line definition, but always shows the same part of the vehicle. In contrast, in a lane change the whole vehicle crosses the scan line in a short amount of time and space. This appears in the epoch as a small spot with the shape of the vehicle.

A Graphical User Interface has been coded to enhance the visual counting from the epochs (see Figure 4). With this GUI, the counting is done by looking at the epoch (bottom of the window) and clicking on the candidate lane change. Then, the video (top of the window) is automatically set at the clicked time in the epoch. If looking at the video frame is not enough, it is possible to play the video around the clicked point to decide if it is a lane change or not. This has the advantage of being more reliable especially for those lane changes happening at the limits of the scanline. The counted lane changes appear in the rightmost table.

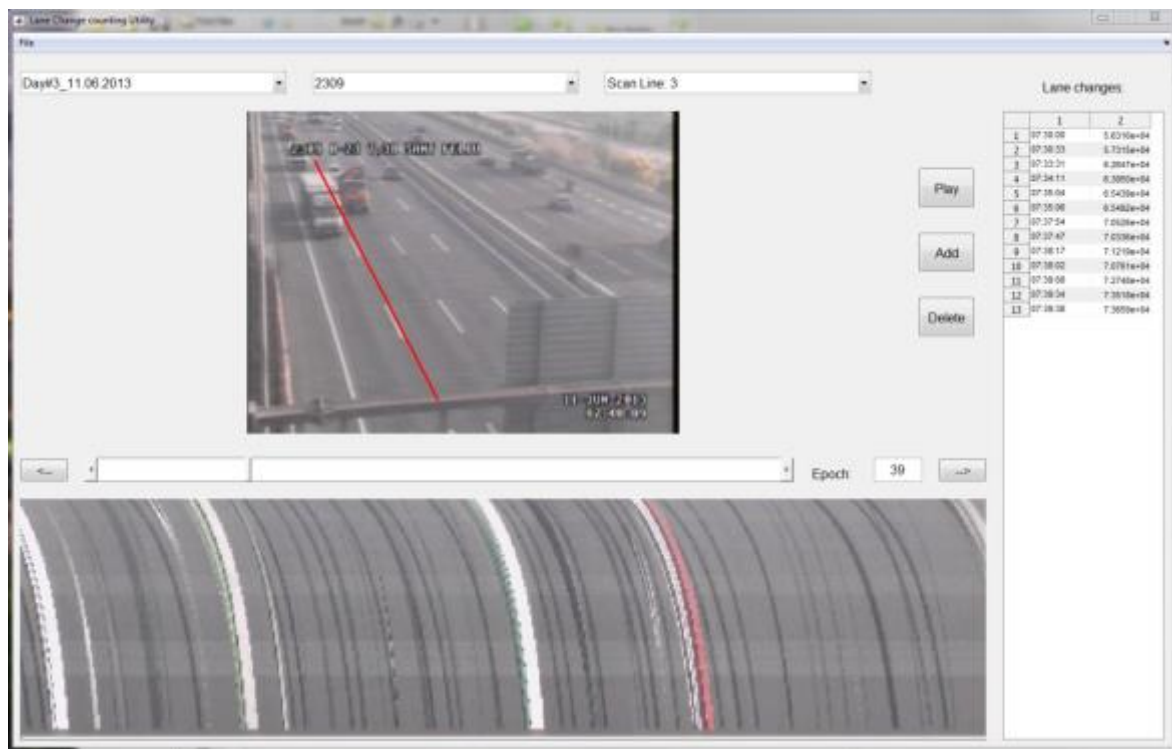


Fig. 4 – GUI for counting lane changes.

3.4 Camera configuration

Cameras must be focused in order to obtain the best possible sharpness. The minimum recommended video resolution is 480x360 pixels, although lower resolutions (e.g. 320x240) can still be used. Regarding the frame rate, 24 fps or higher is recommended, although it is possible to start counting lane changes from 10 fps.

The camera framing is crucial, especially if using low resolution camera settings. The entire image needs to be focused on the freeway. Only "pavement" pixels are useful. The sky or surrounding trees add nothing. Some examples are shown in Figure 5. The visible freeway length must be much longer than the length taken by the typical lane-changing maneuver. To give an approximate value, capturing a freeway length greater than 50 meters is recommended. This allows clearly differentiating occlusions from lane changes. Finally, the frame angle needs to be lined with the line dividing lanes, as much as possible. This minimizes occlusions (see Figure 6).



Fig. 5 - Video framing.



Fig. 6 - Video recording angle

As a consequence of the previous conditions, a correct video framing usually comprises a stretch of freeway starting at least 100 m from the camera site and ending up to 500 m further away.

The technique can also be used to count vehicle flow. In this case, the image requirements are lower. It is only necessary that the vehicles are clearly seen in the video. Then, they will appear in the epochs, and will be counted. However, a frame rate of at least 10 fps is still necessary so that the cars will have a clear vehicle shape in the epoch instead of a blurry stain.

It is important to consider that the better the camera configuration, the shorter the visual processing time. A bad configuration can even invalidate the method, so, always test each camera configuration for a few minutes to see the resulting epoch.

3.5 Environmental recommendations

Not only technical aspects must be taken into account. Environmental factors also have an impact. Clear daylight conditions are necessary. The method cannot be applied to night recordings, because the vehicle headlights dazzle the camera. A similar effect happens during dawn and dusk, so it is recommended not to point the cameras towards the sun. Bad weather, such as fog or rain, can lead to low quality video recordings. However, this may not be a problem, unless the image is so blurry that vehicles cannot be identified.

3.6 Other general considerations

The computer time required for constructing the epochs from videos depends on various factors: video resolution, video frame rate, number of scanlines, length of the scanlines, computational power and video codec used. For the application presented in Section 4, using a Matlab implementation on a computer i-7-4790 CPU with 8GB of RAM memory, it took about half the video duration to convert the video to the epochs. However, the

computer remains fully usable during epoch creation. Thus, it does not consume all the resources available in the computer. It has also been tested in older desktop computers and in a laptop and still remain usable.

4 APPLICATION TO THE B-23 FREEWAY ACCESSING BARCELONA

The methodology developed in this paper has been applied to count the lane changes in some particular locations on the B-23 freeway accessing Barcelona, in the context of a dynamic speed limits (DSL) experiment. See (Soriguera & Sala, 2014) for a complete description of the experiment.

63 hours of standard camera recordings needed to be processed in order to extract lane changing activity. In this context, full automation was infeasible, and the proposed semi-automatic video processing was selected.

The camera framing and the scan line used to create lane change epochs for each pair of lanes in the 6 locations are shown in Figure 7. The video quality consisted on a resolution of 536x400 pixels and a frame rate ranging between 10 and 30 fps. The hardware at the traffic management center (TMC) could not handle higher video resolution.







	Camera framing with lane change epoch lines	Length		Camera framing with lane change epoch lines	Length
Cam 2304		90 m	Cam 2309		70 m 65 m for lanes 1-2
Cam 2305		260 m	Cam 2310		120 m
Cam 2306		115 m	Cam 2312		260 m 250 m for lanes 1-2

Fig. 7 - Lines considered for lane changing counting.

4.1 Performance of the method

The time needed to count the lane changes using the GUI implementation of the semi-automatic method was around 15% of the real video duration. This is more than three times faster than actually watching the video at double speed. In some cameras, with optimal framing and good environmental conditions, lane changes could be directly and accurately extracted from the epochs, without the GUI help. This allowed speeding up the counting up to only 10% of the real video duration.

Nevertheless, by using the GUI the reliability of the lane change counting increases (see Table 1). In such case, the error is small or even zero. In some situations, the usage of the method can even identify more lane changes than actually watching the video. For instance, it has been found that when two lane changes happen at different points but almost at the same time, only one is seen from the video. In the epoch, both can be clearly identified. These additional lane changes have been confirmed by playing the video at slower speeds.

Camera	2304	2305	2306	2309	2310	2312
Lanes	2-3	1-2	1-2	3-4	2-3	1-2
Day	1	5	6	2	3	4
Duration of this accuracy test (10 min)	07:10 to 07:20	08:00 to 08:10	09:00 to 09:10	7:30 to 7:40	08:30 to 08:40	09:00 to 09:10
Video count	16	22	21	12	19	21
Epoch count	14	21	20	3	4	6
Epoch + GUI count	16	28*	21	12	19	13

Table 1- Counting lane changes with different methods (*Additional lane changes counted with GUI. The error in this case is in the manual counting).

By only using the epochs, without the GUI, the results are acceptable in cameras 2304, 2305 and 2306 which have good video recordings. Assuming the manual count as ground truth, the average relative error in these cameras is 6,7%. However, cameras 2309, 2310 and 2312 have errors above 50%. These are greatly improved by using the GUI. Still, bad results are obtained in camera 2312. This is due to the bad camera framing, leading to a high rate of occlusions, especially for lanes 1-2.

5 SOME PRELIMINARY RESULTS

A preliminary analysis of the resulting lane changing database, allowed to empirically prove that lane changing peaks at traffic state transitions (i.e. free flowing => congestion, or vice versa). This can be clearly seen in Figure 8, where cumulative count (N-curve), cumulative occupancy (T-curve), and cumulative lane changes (L-curve) with respect to time are plotted. The transition between free flow to congestion is identified by an increase of the occupancy (slope of the T-curve) and a decrease of the flow (slope of the N-curve). This happens slightly before 8:30 in Figure 8. Free flow recovering is identified by an occupancy drop (i.e. around 9:00 in Figure 8). In both situations, the lane changing rate peaks (slope of the L-Curve). Although further research is needed, this only pretends to show an interesting research direction that the presented method supports.

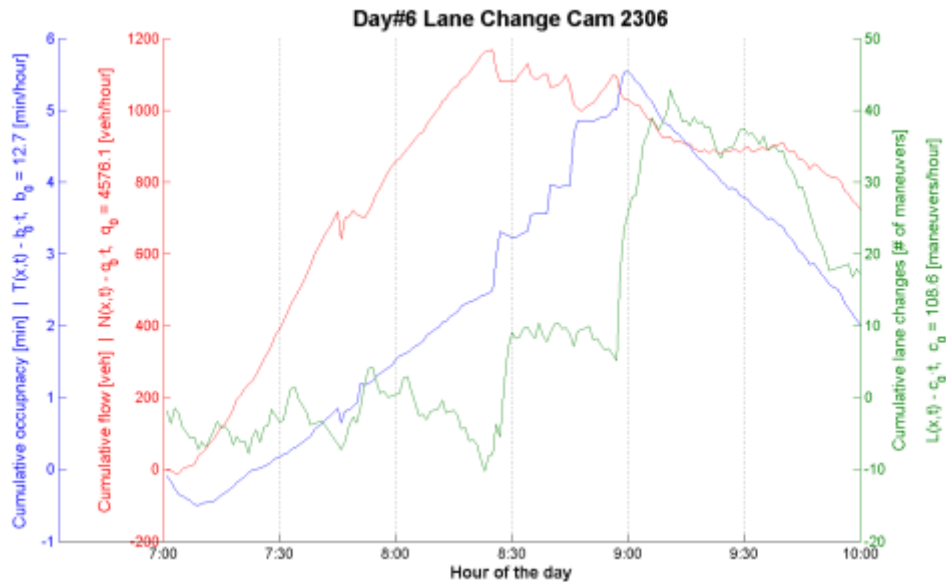


Fig. 5 - Oblique cumulative count (N), occupancy (T) and lane change (L) curves.
 Note 1) Data is obtained from camera 2306 on wed. 13rd June 2013 (Day#6). 2) See (Cassidy & Windover, 1995) for a description of the cumulative count curve methodology. The subtracted background flow is a 95% of the average.

6 CONCLUSIONS

Manually extracting traffic data from video recordings requires multiple video visualizations, it is time consuming and extremely tedious. For relatively short applications, full automatic methods are overcomplicated. The method presented in this paper aims to fill this gap. It develops a semi-automatic and simple video processing tool, which eases and speeds up the manual process without falling in the complexities of complete automation.

The method is based on the transformation of the video recordings to a set of adequate images (i.e. called "epochs"), where lane changes (or vehicle counts) are visually identified, easily and fast. The method has been implemented on a user friendly GUI. From the application to 68 hours of recording on a Spanish freeway in order to extract lane changing movements, it is estimated that the method reduced total human observation time to 10-15% of the total video duration, without any significant reduction in the accuracy. In some cases, (e.g. when two lane changes happen at the same time) the proposed method is even more reliable than actually watching the video.

In order to successfully use the method, it is crucial to follow the recommendations regarding the video quality and framing. The better the video quality is, the easier, faster and more reliable the counting.

7 ACKNOWLEDGEMENTS

Authors acknowledge the collaboration of the *Servei Català de Trànsit*, the Catalan traffic administration for providing the required video recordings. We are especially grateful to Carles Argüelles and Javi Romero and their teams (*UTE acc. ACISA-sur Aluvisa*) who

were in the TMC during the experiment and dealt with a significant amount of extra work recording all the data. This research has been partially funded by the Spanish Ministry of Science and Innovation (TRA2013-45250-R/CARRIL).

8 REFERENCES

- Barragán, D., 2015. Traffic surveillance using image processing. Available at: <http://es.mathworks.com/matlabcentral/fileexchange/49554-traffic-surveillance-using-image-processing> [Accessed March 1, 2016].
- Bhaskar, P.K. & Yong, S.-P., 2014. Image processing based vehicle detection and tracking method. *2014 International Conference on Computer and Information Sciences (ICCOINS)*, pp.1–5. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6868357>.
- Campbell, R.J., 2012. *An Analysis Framework for Evaluation of Traffic Compliance Measures*. University of California at Berkeley.
- Campbell, R.J. & Skabardonis, A., 2013. Analysis framework for evaluation of traffic compliance measures. *Transportation Research Record*, (2364), pp.71–79. Available at: <http://dx.doi.org/10.3141/2364-09>.
- Cassidy, M.J. & Windover, J.R., 1995. Methodology for assessing dynamics of freeway traffic flow. *Transportation Research Record*, (1484), pp.73–79.
- Federal Highway Administration, 2006. *Interstate 80 Freeway Dataset Factsheet*, Available at: <http://www.fhwa.dot.gov/publications/research/operations/06137/>.
- Federal Highway Administration, 2015. Next Generation Simulation (NGSIM). Available at: <http://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm> [Accessed March 3, 2016].
- Lu, X. & Skabardonis, A., 2007. Freeway Traffic Shockwave Analysis: Exploring NGSIM Trajectory Data. *Transportation Research Board 86th Annual Meeting*, (January 2007), p.19.
- Oliveira, M. & Santos, V., Automatic Detection of Cars in Real Roads using Haar-like Features. , pp.1–6.
- Patire, A.D., 2010. *Observations of lane changing patterns on an uphill expressway*. University of California, Berkeley.
- Patire, A.D. & Cassidy, M.J., 2011. Lane changing patterns of bane and benefit: Observations of an uphill expressway. *Transportation Research Part B: Methodological*, 45(4), pp.656–666. Available at: <http://dx.doi.org/10.1016/j.trb.2011.01.003>.
- Punzo, V., Borzacchiello, M.T. & Ciuffo, B., 2011. On the assessment of vehicle trajectory data accuracy and application to the Next Generation Simulation (NGSIM) program data. *Transportation Research Part C: Emerging Technologies*, 19(6), pp.1243–1262. Available at: <http://dx.doi.org/10.1016/j.trc.2010.12.007>.
- Soriguera, F. & Sala, M., 2014. Experimenting with Dynamic Speed Limits on Freeways. *Procedia - Social and Behavioral Sciences*, 160(CIT), pp.35–44. Available at: <http://dx.doi.org/10.1016/j.sbspro.2014.12.114>.
- Uke, N.J. & Thool, R.C., 2013. Moving Vehicle Detection for Measuring Traffic Count Using OpenCV. *Journal of Automation and Control Engineering*, (FEBRUARY), pp.349–352. Available at: <http://www.joace.org/index.php?m=content&c=index&a=show&catid=35&id=151>.