

Automatic Traffic analysis in video sequences

Kannikar Intawong*, Mihaela Scuturici* and Serge Miguet*

Abstract—The main goal of this study is to extract traffic statistics throughout the day, week and season based on video surveillance. For our propose, we focus on traffic analysis that is related to vehicle counting. We provide the following process chain: objects detection and tracking, trajectories extraction, detection of entry/exit zones and vehicle counting. Experiments show that the information obtained is reliable and gives effective vehicles counting that are compared to previously published results obtained on a public data set. Additional comparisons are done with ground truth for a private data set.

Keywords—image processing, video surveillance, traffic analysis, statistics

I. Introduction

For several years, video surveillance has developed quickly in many application contexts, such as traffic monitoring or autonomous vehicle guidance [1], [2] supervision of elderly or sick persons [3] and shopping or retail fraud detection [4]. In a general way, the goals of a video surveillance system are the following: detection of a specific object, extraction of statistical information (such as for example, the average number of cars, trucks, bicycles), or the detection of specific events (stopped vehicles, accidents etc.).

In this paper we define video analytics as computer-vision-based surveillance algorithms and systems to extract contextual information from video. The main concept is to aid human operators in observing video data. Video cameras have been deployed for a long time for traffic and other monitoring purposes, because they provide a rich information source for human understanding. Video analytics may now provide added value to cameras by automatically extracting relevant information. This way, computer vision and video analytics become increasingly important for intelligent transport systems (ITSs). The method proposed in this paper is related to automatic counting. To be useful, a video surveillance system must be fully automatic and able to provide, in real time, information that concerns the object's behaviors in the scene. There are several approaches that aim at automatic traffic analyzing. But most of them have difficulties to handle particular situations. Nowadays, there are many challenges such as occlusion, long-time vehicles stops and illumination changes, etc.

Our processing chain is composed of the following steps:

The detection of moving objects is based on a parametric version of the bloc-based hierarchical mixture of gaussians [5] which is considered in [6] as one of the best state of the art method for object extraction. We have shown in [12] a methodology used to choose the optimal value of the parameters of this algorithm.

Object categorization and vehicle extraction using fourier descriptors (pedestrian, bicycle, animals must be filtered out).

Object tracking, using a kalman filter and a descriptor based on dominant colors. Because of above mentioned difficulties, trajectories we obtain might be incomplete, resulting to wrong counting information.

In order to make these systems more robust for individual tracking, we propose to aggregate the partial data to build global information on the flow of vehicles in the scene. Specifically, we propose a method which, automatically identifies input E_i and output X_j areas in the scene and we record the number of vehicles for each pair E_i, X_j .

II. Related work

This section reviews applications and existing commercial systems for traffic monitoring. Our approach will focus on vehicle counting, which is mainly applied to highway scenes. The problem of vehicle counting is most commonly solved by deploying inductive loops. These loops provide high precision but are very intrusive to the road pavement and, therefore, come with a high maintenance cost. Most video analytics systems on highways focus on counting and, possibly, classification to allow for more detailed statistics [1], [3], [4]. Previous works relate some problem that lead to counting areas: wrong split of objects due to occlusions [8] or incorrect segmentation due to sudden illumination changes [2]. Our approach will try to solve at a global level most of the local errors that are reported by previous.

III. Methodology

Firstly, we extract objects by using background modeling and track the objects. The output of the tracking process is a set of trajectories, it begins with incomplete trajectories (also known as partial flow). Second is the aggregation of partial flow in order to build a global flow. We divide the image into disjoint cells and collect the directions of moving objects. We propose to use the regular grid structure and local information between 8 neighboring cells. This data structure called "matrix of motion direction" has the advantage of building a robust scene that can help the tracking process by giving a reliable motion prediction representation. On blocks of 16x16 pixels, we calculate a map of movement directions, where the

* Université de Lyon, CNRS
Université Lumière Lyon 2, LIRIS UMR5205
5, Av. Pierre Mendès-France, 69676, Bron, France

directions of the trajectories of moving objects are coming out of each block. From the global flows, we extract the exit points of vehicles by direct scattering and entry points by reverse scattering. Thirdly, we find the entry and exit zones E_i and X_j of entire scene by using points clustering from Global reverse flow and Global direct flow respectively. Finally, we calculate automatically the statistics of the moving vehicles in trajectories

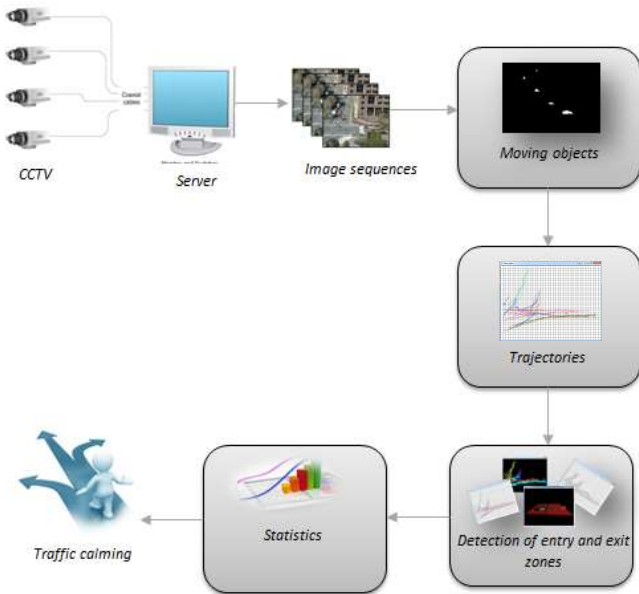


Figure 1. Traffic analysis system overview.

A. Object detection

In order to segment moving objects we use the hierarchical background modeling technique introduced by [5]. It is considered as one of the best background modeling algorithms in the comparative study proposed by [6]. This method uses coarse level contrast descriptors whose evolution is represented by Mixtures of Gaussians. This coarse-level representation is then combined with the classical pixel-based Mixture of Gaussians. We can thus identify the foreground objects at coarse level, then detail shapes of foreground objects at pixel level. Figure 2 presents some results of background subtraction algorithms.

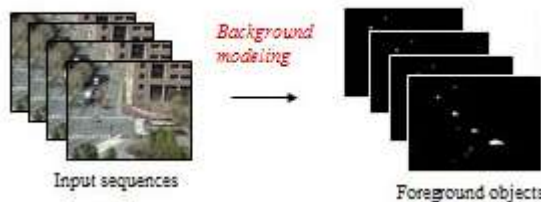


Figure 2. Object detection.

Object detection by background modeling algorithms, used without postprocessing, very often let appear isolated pixels in the background. They are considered as foreground objects. On the contrary, holes in objects are classified as background. The method most commonly used to overcome these drawbacks is to apply mathematical morphology (erosion, dilation, opening, closing, . . .).

In our experiments we vary the number of erosions and dilations applied as post-processing of the segmentation. We search the parameter sets giving the best possible values for each of the segmentation quality measures.



Figure 3. Left is before morphological operations and right is after morphological operations.

B. Evaluation measure

In order to improve segmentation, we use quality measure to find good values for input parameters of image segmentation algorithms.

The results of any segmentation algorithms vary as a function of the values of different parameters. The best parameters values minimize a distance or maximize the similarity between the segmentation and the ground truth. The F-measure, the Percentage of Correct Classification (PCC) and the Jaccard coefficient (JC) are some of the measures, we can try to optimize in order to have the best possible segmentation.

Another approach[12] that we use in this paper which to measures the quality of a segmentation as an Euclidean distance called D_{prs} in the space of the indicators, between the point (PR, RE, SP) and the ideal point (1, 1, 1). This can be seen as a 3D generalization of the use of 2D ROC curves or Precision/Recall diagrams for optimal parameters determination.

We have shown in [12] that our approach for determining the optimal parameters gives better results than traditional approaches.

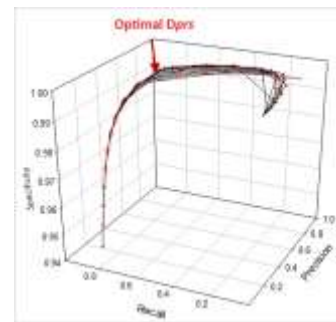


Figure 4. Precision, Recall, Specificity space.

C. Trajectories

A foreground object is characterized by its boundary, position centroid and color. A Kalman filter which models the position, size and velocity of each blob assists the matching process. Then, we use Fourier descriptors to classify the objects as vehicles and pedestrians, because we are interested in vehicle counting only. The output of the tracking process is a set of trajectories. A trajectory aims to describe the location history of a target moving through the scene. Information about the position, size, and color of the corresponding blob is held for each frame over which the object was tracked.

D. Detection of entry and exit zones

- Building the matrix of motion directions

For each 16x16 image block, we consider the outgoing directions of moving objects (see Figure 5). We count how many vehicles leave the current block towards the 8 neighbors. Assume the input frames are of size $h \times v$.

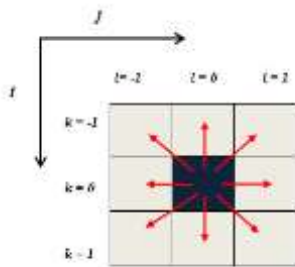


Figure 5. Motion directions of each block.

- Weight evolution

The main idea of our algorithm for exit point computation is based on an iterative process that makes this weight matrix evolve, following the directions stored in the matrix of motion directions.

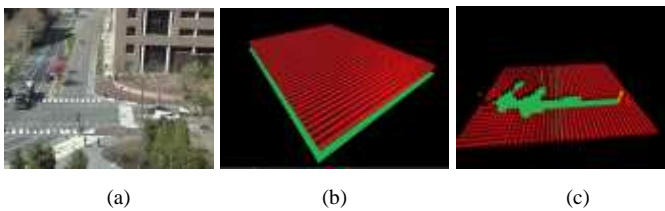


Figure 6. Weight evolution concept: (a) Original images, (b) weight matrix initialization. Weight evolution: cells traversed by moving objects are emptied, (c) weights are transferred towards exit points (in pink), where they accumulate.

The matrix of motion directions is then normalized using pseudo code of algorithm 1. After this normalization step, $s[i, j, k, l]$ contains the proportion of vehicles that leave cell $C[i, j]$ in direction of cell $C[i+k, j+l]$.

```

Algorithm 1: Matrix normalization
1 sum = 0
2 for i ← from 0 to v-1 do
3   for j ← from 0 to h-1 do
4     for k ← from -1 to 1 do
5       for l ← from -1 to 1 do
6         if k ≠ 0 and l ≠ 0 then
7           s[i, j, k, l] = sum
    
```

The content of each cell is scattered between the eight neighboring cells, proportionally to the values stored in the matrix of motion directions. By iterating this process following algorithm 2, the weights will accumulate in cells having only incoming motion.

```

Algorithm 2: Weight matrix iteration
Input:  $W_t$ 
Output:  $W_{t+1}$ 
1  $W_0 \leftarrow [1]$ 
2 for  $t \leftarrow 1, 2, \dots, n$ 
3 do
4    $W_{t+1} \leftarrow W_t$ 
5   for  $i \leftarrow from 0 to v-1 do$ 
6     for  $j \leftarrow from 0 to h-1 do$ 
7       if  $sum[i, j, k, l] \neq 0$  then
8         for  $k \leftarrow from -1 to 1 do$ 
9           for  $l \leftarrow from -1 to 1 do$ 
10            if  $k \neq 0$  and  $l \neq 0$  then
11              if  $s[i, j, k, l] \neq 0$  then
12                 $\alpha = \frac{W_t[i, j] * s[i, j, k, l]}{\sum_{k=-1, l=-1}^1 s[i, j, k, l]}$ 
13                 $W_{t+1}[i+k, j+l] += \alpha$ 
14                 $W_{t+1}[i, j] -= \alpha$ 
15 return  $W_{t+1}$ 
    
```

- Construction of entry/exit zones

The different exit zones are obtained by clustering exit points, while taking into account the average apparent size of the objects in each cell: far objects are small and exit points need to be closer to each other in order to be considered in the same exit zone. In the same manner, near objects are larger, allowing the construction of larger exit zones (see Figure 7). Entry points and entry zones are obtained by the same algorithm, by reversing the direction of the trajectories.

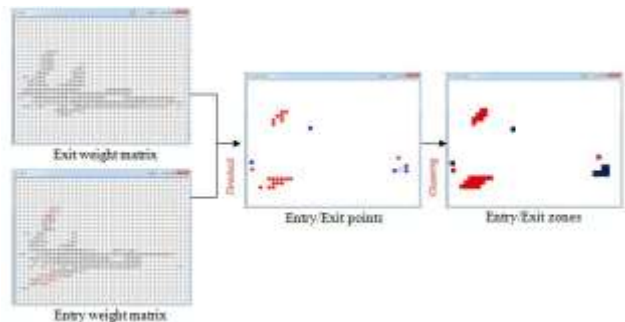


Figure 7. Entry and exit zones detected by our approach.

E. Counting strategies

The two strategies we propose differ depending on whether a partial trajectory might be merged with another only once (strategy1) or possibly to several other partial trajectories (strategy2).

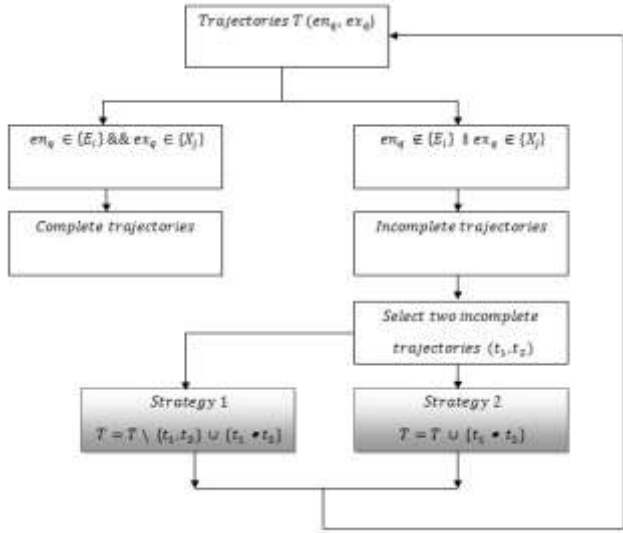


Figure 8. Counting strategies.

IV. Experimental results

- Entry /Exit zones detection

Our experiments has been computed for six video sets composed of 3000 frames each which are Chiang Mai traffic recording, video of USA highway[11] and video of MIT traffic dataset[7]. Figure 9 is a step by step illustration the detection of Entry/Exit zones of the results we obtained in our experimental work explained in processing chain.

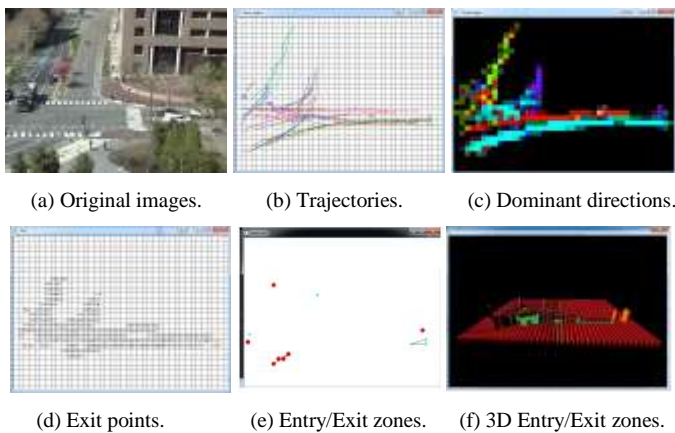


Figure 9. Experimental results.

- Comparison

We compare the counting results of our algorithms to the ground truth obtained by a manual counting. For MIT dataset we compare our approach to the results of a previously published paper [8]. These experiments seem to show the ability of our approach to handle complex situations which could not be processed by the approach presented in [8] (see figure 10(b)). A particular situation is illustrated in Figure 5: due to multiple lane crossings, the semantic regions obtained in [8] give incorrect entry/exit zones in the scene. Our approach succeeds in determining correct zones and is more precise in automatic statistical analysis of vehicles flow.

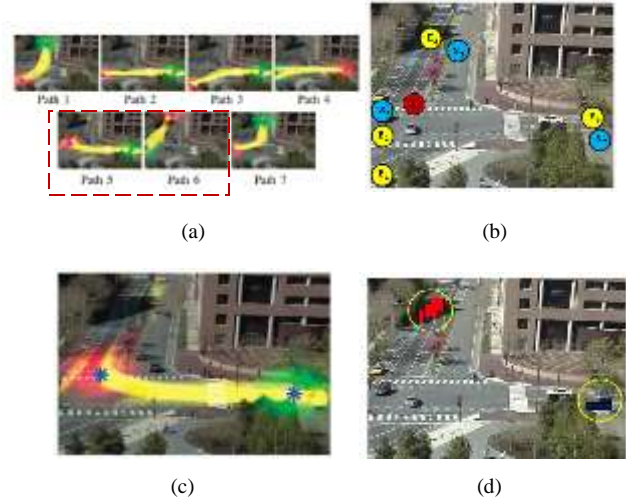


Figure 10. The comparison results: (a)The seven paths detected in [8] (b) Seven paths in [8] matching to our entry/exit zones. The source of Path 5 does not exist in our approach because it is not a real entry zone, (c) Example of error of path 5 in [8] (d) Entry/Exit zones by our approach.

- Statistics

To evaluate the performance of the traffic counting, we use the precision and recall [9] by considering the assignment of each vehicle to the correct (E_i, X_j) pair. We assume one complete trajectory is equivalent to one vehicle for counting.

Table I shows global results obtained on MIT data set by using our method. In this video, strategy 2 gives more accurate results than strategy 1.

TABLE I. COMPARISON OF STRATEGY 1 AND STRATEGY 2 RESULTS ON THE MIT SCENE.

Method	Ground truth	Results	Recall	Precision	F-measure
Before completion	44	27	0.42	0.60	0.494
Strategy 1	44	52	0.90	0.84	0.868
Strategy 2	44	47	0.82	0.94	0.875

We also calculate statistics on 6 different sequences of Chiang Mai video data set, USA highway video and MIT dataset. Our approach gives correct entry/exit zones in the scene (see figure 11). For the six sequences, strategy 2 gives

more effective than strategy 1. We obtain good precision values around 91% and F-measure averages is 0.90 for vehicle counting as shown in Table II. However, this results show the impressive of the proposed technique.

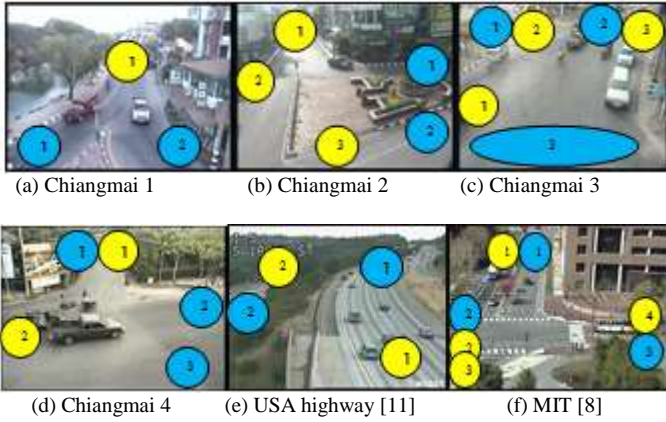


Figure 11. Entry zones(yellow circles) and Exit zones (blue circles).

TABLE II. COMPARISON OF STRATEGY 1 AND STRATEGY 2 IN GLOBAL VIDEOS.

Video	Ground truth	Strategy 1				Strategy 2			
		Out put	RE	PR	F	Out put	RE	PR	F
Chiangmai1	63	57	0.90	1	0.94	65	0.98	0.96	0.96
Chiangmai2	21	27	0.92	0.77	0.83	24	0.90	0.85	0.90
Chiangmai3	18	25	0.90	0.72	0.80	22	0.88	0.81	0.84
Chiangmai4	52	61	0.94	0.85	0.89	55	0.90	0.90	0.90
USA highway	34	29	0.85	1	0.92	29	0.85	1	0.92
MIT	44	52	0.90	0.84	0.87	47	0.82	0.94	0.88

In Table III, we compare the results between Bayesian approach, semantic regions introduced in [8] and our method using strategy 2 for trajectory completion. In the ground truth there are 44 vehicles but Bayesian algorithm collects 59 vehicles and semantic region method finds 48 vehicles. In our approach, one completed trajectory represents one vehicle and we find 47 vehicles. The precision of semantic regions is 0.92, of our approach is 0.94 and the precision of Bayesian is only 0.74.

TABLE III. COMPARISON OF COMPLETED TRAJECTORIES RESULTS BETWEEN BAYESIAN, SEMANTIC REGIONS AND OUR APPROACH COUNTING ON MIT SCENE.

Method	Ground truth	Results	False alarms	Precision	F measure
Bayesian [Zhao and Wang]	44	59	15	0.74	0.850
Semantic regions [Zhao and Wang]	44	48	4	0.92	0.958
Our approach (strategy 2)	44	47	3	0.94	0.970

V. Conclusion

This paper presents a complete traffic video analysis system. We propose the process chain that is related to traffic counting: detection and tracking, trajectory, Entry/Exit zones and Statistic. Experimental results demonstrate the efficiency of our algorithm for improving the quality of automatic system. This algorithm allows us to obtain a very reliable analysis of traffic video. The perspectives of this work will be count the number of pedestrian and analyze the real-time video.

Acknowledgment

This work is partly supported by Control Traffic Center from Chiang Mai Municipality, Thailand and LIRIS laboratory (d'Informatique en Image et Systèmes d'information), UMR 5205 CNRS, Université Lumière Lyon 2, France.

References

- [1] N. Buch, S.A. Velastin, J. Orwell, A review of computer vision techniques for the analysis of urban traffic, IEEE Transact. Intell. Transp. Syst. 12 (3) , 2011, 920–939.
- [2] H. Nicolas, M. Brulin, Video traffic analysis using scene and vehicle models. in: Signal Processing: Image Communication, 2014, pp. 807–830.
- [3] S. Karaman, J. Benois-Pineau, V. Dovgalecs, R. Mégret, J. Pionquier, R. André-Obrecht, Y. Gaestel and J.F. Dartigues. Hierarchical Hidden Markov Model in Detecting Activities of Daily Living in Wearable Videos for studies of Dementia. Arxiv preprint (arXiv:1111.1817), 2011.
- [4] H. Trinh, Q. Fan, P. Jiyang, P. Gabbur, S. Miyawaza, S. Pankanti, Detecting human activities in retail surveillance using hierarchical finite state machine. in: Proceedings of ICASSP, IEEE, 2011, pp. 1337– 1340. Writer's Handbook. Mill Valley, CA: University Science, 1989.
- [5] Chen, Y.T., Chen, C.S., Huang, C.R., Hung, Y.P.: Efficient hierarchical method for background subtraction. Pattern Recognition, 2007, pp.2706-2715.
- [6] Dhome , Y., Tronson, N., Vacavant, A.: A Benchmark for Background Subtraction Algorithms in Monocular Vision: a Comparative Study. Image Processing Theory Tools and Applications (IPTA) , 2010.
- [7] MIT. Traffic dataset.. URL : <http://www.ee.cuhk.edu.hk/xgwang/MITtraffic.html>, 2009.
- [8] R. Zhao and X. Wang. Counting vehicles from semantic regions. In Intelligent Transportation Systems, 2013, pp. 1016–1022.
- [9] J.Davis and M. Goadrich. The relationship between precision-recall and roc curves. In International conference on Machine learning, 2006, pp. 233–240..
- [10] I. Creusen, R.Wijnhoven, and P. H. N. deWith, “Applying feature selection techniques for visual dictionary creation in object classification,” in Proc. Int. Conf. IPCV Pattern Recog., Jul. 2009, pp. 722–727.
- [11] USA Highway data set URL: <http://opendata.stackexchange.com/questions/1767/data-of-vehicle-traffic>
- [12] K. Intawong, M. Scuturici, S. Miguet, A New Pixel-Based Quality Measure for Segmentation Algorithms Integrating Precision, Recall and Specificity,International, Conference on Computer Analysis of Images and Patterns, 2013, pp. 188-195.
- [13] CRS, Computer Recognition Systems. [Online]. Available: <http://www.crs-traffic.co.uk/>.