Vehicles Detection From Aerial Sequences

Khaled Kaâniche, Pascal Vasseur, and Claude Pégard

Université de Picardie Jules Verne, Centre de Robotique d'Electrotechnique et d'Automatique,
EA 3299 - 7 rue du moulin neuf 80000 Amiens

khaled.kaaniche@u-picardie.fr
Pascal.Vasseur / claude.pegard@sc.u-picardie.fr

Our work is based on aerial sequences taken from an UAV-Camera system that we use to localize one or several targets (in our case : vehicles). The dynamic behavior of the UAV-Camera system make having fixed background impossible. So, scene processing becomes more complicated. Also, this behavior generates noise due to vibrations. In this paper, we propose two approaches which aim to detect vehicles from aerial sequences.

1 Introduction

In this work, we propose two approaches which aim to detect vehicles from aerial squences. The first approach is based on static analysis of the sequence (frame by frame), the second is based on the observation of the primitives motion in time. Our purpose is to create, first of all, a graph. Nodes of this graph are the primitives detected in one frame by segmentation. Those nodes are interconnected by two types of links : in the first approach links are based on perceptual organization criteria -mainly geometric- ; in the second approach links consider every node motion and their configurations in the next frame. In other words, grouping through time is based on "common fate" of primitives from frame to frame. This graph is used as the root of all the grouping process. It's on it that the partition will be executed. For this step of partition, we choose "normalized cuts technics" [13]. Shi and Malik [13] consider visual grouping as a graph partitioning problem in which nodes are simply the pixels and links the degree of similarity in intensity or in color of those pixels. In our study, primitives detected are used as nodes. Obviously, the number of primitives is less important than the number of pixels in the original image. Through this choice, we relieve the matrix computation. However, this choice may cause a lose of information as an image has more information than its binary derivative. We justify this risk by the fact that the proposed grouping process relies mainly on the geometric aspect of the primitives shown in the edges.

1.1 Previous Work

Extracting roads and vehicles from aerial sequences took from an helicopter (dynamic behavior of the view point) is difficult. In fact, this type of applications has been rarely treated in the literature. Most study deals with static images. Burlina, Parameswaran and Chellappa [2], Liu and Haralick [8] and Moon, Chellappa and Rosenfeld [11] consider vehicles as 2D elements, a vehicle is modelled as a rectangle. The model is socked in the binary frame ; problem is then considered as an edge detection. Zhao and Nevatia [16] attempted to improve this process by formulating the problem in 3D, the error on the final decision is smaller as the model has higher information content. Both quoted methods define a vehicle model (2D or 3D) and look for it in the image. In the contrast, we don't define any model, we just let geometric and motion criterion decide of the way the sequences parts would be grouped. Road traffic analysis was the object of a project done by Malik and Russell [9]. The main difference between this project and our application is the dynamic behavior of our capture tool. With a static capture tool, a simple subtraction between the acquired pictures and the updated background is sufficient to detect what moves in the scene (vehicles in this case). Medioni and Cohen [10] propose a system able to detect moving object in aerial sequences. The idea consists on the estimation and the compensation of the camera motion. Moving object correspond to the regions with residual motion. We treat the problem differently, we are interested to the perceptive aspect in the sequences.

The organization of this paper is as follows. In section 2 we present the first approach based on perceptual organization criteria. Section 2 shows also some results using this method. In the section 3, we develop the second approach based on "common fate" analysis and we present some results obtained thanks to this approach.

2 Vehicles detection based on perceptual criteria

This approach [6] is divided on three parts 1: the segmentation of the acquired images, the modelization of a graph in which primitives provided by the segmentation are connecting according to perceptual criteria [1] (Proximity and Parallelism) , and the bi-partition of the graph by normalized cuts technique [13]. The final aim consists on extracting vehicles from the background. Parameters of the proposed algorithm are chosen after a learning stage in which we maximize the similarity between manual cut and normalized cut results. Genetic Algorithm are used in this step of optimization [7].

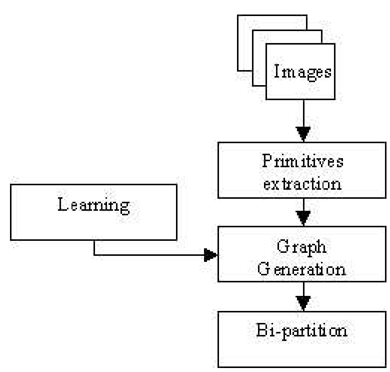


Fig. 1. Vehicles Detection based on Perceptual Criteria.

2.1 Perceptual Organization

Sarkar and Boyer [1] define the Perceptual Organization as the ability of a vision system to organize detected features or primitives in image based on Gestaltic criteria for example. Gestalt psychologists have offered a set of laws that are important in figure-ground segmentation like the laws of parallelism, continuity, proximity, similarity, common region and symmetry [12]. We will use this type of laws to set connections between the nodes of our graph ; thus, we will refer to a simple process based on calculating a score (or a probability) between primitives for a specific law. For example, if we choose proximity as a grouping criteria, the process computes a score nearing *1* if the two primitives touch each other. This score becomes *0* as soon as the distance between the two primitives exceeds a fixed threshold. Then, the score must be high to favor the grouping of two nodes according one law. However, using several laws or criterion at the same time seems to be important for complex applications. One law can not extricate by itself various types of shapes or entities. So, combining many scores becomes necessary. This can be done by logic operators or level-headedness depending on applications and experimental tests. Finally, the choice of laws or criterion can be justified by the the kind of application. That's why, in the specific context of road traffic, where we look for detecting roads and vehicles, we believe that parallelism and proximity are among suitable laws.

2.2 Normalized Cuts

As one of the new techniques conceived for the partitioning of graph, Normalized Cuts [13] is a method based on the minimization of a similarity criterion connecting nodes of different parts. We should also mention Minimum Cuts [15] and Average Cuts [12]. All those partitioning graph processes were developed in a comparative study of Soundararajan and Sarkar [14]. Minimum Cuts process showed limits. In fact, Wu and Leahy [15] like Shi and Malik [13] noticed that the minimum cut criteria favors cutting small sets of isolated nodes in the graph. Both other technics present similar performances in term of correct results. In an explicit way, partitioning graph problem treated by Normalized Cuts is formulated as follows :

Take *G(V,E)*, we want to find A and B / $A \bigcup B = V$ and $A \bigcap B = \emptyset$ which minimize :

$$Ncut(A,B) = \frac{cut(A,B)}{Assoc(A,V)} + \frac{cut(A,B)}{Assoc(B,V)} \tag{1}$$

Where :

$$cut(A,B) = \sum w(u,v) \; ; u \in A \text{ and } v \in B$$
$$Assoc(A,V) = \sum w(u,v) \; ; u \in A \text{ and } v \in V$$

w(u,v) is the weight on the edge connecting the nodes u and v. The second eigenvector of the generalized eigenvalue problem :

$$Wx = \lambda Dx \tag{2}$$

approximates the optimal partition of *G(V,E)* where :

$$W(i,j) = w_{ij} \quad and \quad D(i,i) = \sum w_{ij}, j \in V \tag{3}$$

w_{ij} is the score or probability computed by the PO process for two primitives *i* and *j*. When we use primitives that result from segmentation as nodes, the dimension of the matrix *W* is smaller, so the resolution of the system 2 is almost immediate. This fact is very important because in our application we treat video scenes and consequently, a fast succession of a great number of images.

2.3 Approach

Our algorithm will be applied to a road sequences acquired from UAV-camera system. Sequences are cut in successive fixed images. System block diagram described in Fig. 1, shows different stages of our algorithm. Firstly, a Canny-Edge-Detector [3] is used to extract primitives from successive images. This detector allows also to have important information about every primitive like dimension, start-point and end-point coordinates, start and end slope ...etc. We use these information for the computation of weights connecting different nodes or primitives. If we consider *n* as being the number of primitives (nodes) existing in one frame, the size of the matrix *W* will be [$n \times n$] and also for the matrix *D*. The second part of our algorithm consists in calculating these matrices. The matrix *W* groups the weights w_{ij}. In our application the parallelism and proximity criteria are very significant because we treat a road traffic, so we choose to compute w_{ij} with these criteria. w_{ij} is formulated as follows:

$$w_{ij} = Spar_{ij} \times Sprox_{ij} \tag{4}$$

where :

- $Spar_{ij}$ is the score awarded to the degrees of parallelism between two primitives *i* and *j*.
- $Sprox_{ij}$ is the score awarded to the degrees of proximity between *i* and *j*.

$Spar_{ij}$ and $Sprox_{ij}$ are computing as follows :

$$Spar_{ij} = \begin{cases} 1 - \frac{|a_i - a_j|}{seuilpar} & if \quad |a_i - a_j| < seuilpar \\ 0 & otherwise \end{cases} \tag{5}$$

$$Sprox_{ij} = \begin{cases} 1 - \frac{min(d_r)}{seuilprox} & if \quad min(d_r) < seuilprox \\ 0 & otherwise \end{cases} \tag{6}$$

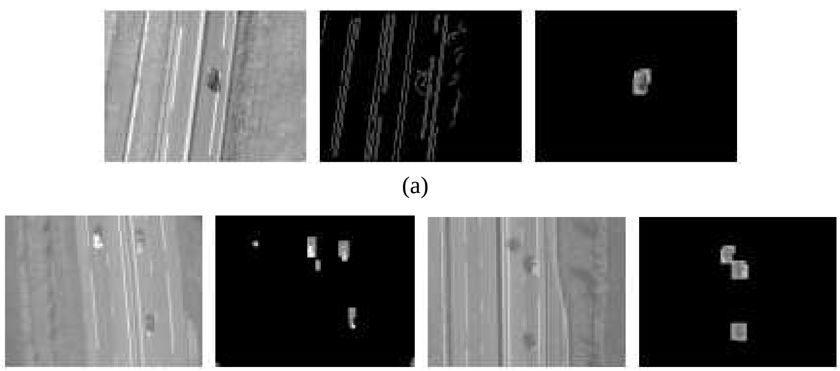
where :

- a_i and a_j are respectively the orientations of the nodes *i* and *j*.
- *min*(d_r) is the smallest distance between nodes *i* and *j*.
- *seuilpar* and *seuilprox* are predefined thresholds.

2.4 Experimental results

In this part, we expose some results done in open-loop environment. Images size is [*300* $\times$ *400*]. Thresholds are inspired from the learning phase developed in [7]. The necessary time for all the operation (without Canny segmentation) is lower than one second. Coding tool is Matlab. Sequences acquirement conditions are the following :

	Fig. 2(a)	Fig. 2(b)	Fig. 2(c)
altitude (m)	415	413	452
latitude	4631'N	4635'N	4633'N
longitude	319'E	323'E	323'E
thresholds	[.1 50]	[.05 50]	[.02 100]



(a)

(b) (c)

Fig. 2. (a) Sequence presenting one vehicle. (b) Sequence presenting two flow directions. (c) Sequence presenting several vehicles.

3 Vehicles detection using motion criterion

Fig. 3 shows the different steps used to extract vehicles from sequences. In general, vehicles motions are different from background motion (caused by the UAV-Camera system move). We use corners [4] data to recognize primitives into successive images and normalized cuts techniques [13] to extract homogenous edges (edges having the same move). The grouping process is then based on "common fate" of primitives from frame to frame.

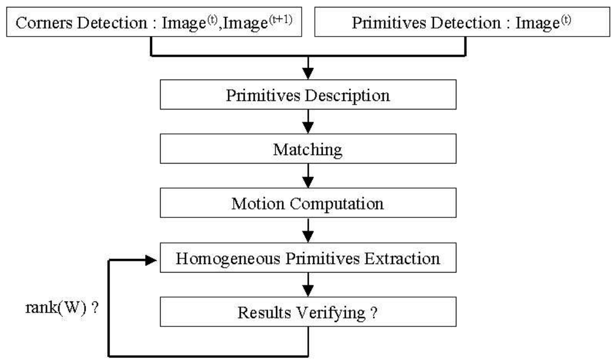
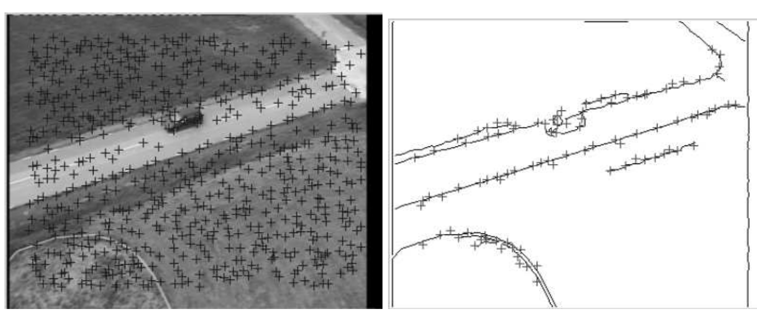


Fig. 3. Vehicles Detection based on motion Criterion.

3.1 Approach

Fig. 4(a) shows an example of corners data detected on one aerial image using Harris [5] detector. We use these corners to describe edges or primitives in the image. every primitive is described by the nearest corners (Fig. 4(b)). The second step is the matching process. Take a set of corners which describe one primitive, matching process consists to find for every corner its corresponding in the next image. Thus, we can compute the mean displacement of the set of corners which describe one primitive and suppose that is the displacement of the primitive between two successive images. Links between nodes (primitives) of the graph is based on the difference between their displacements.

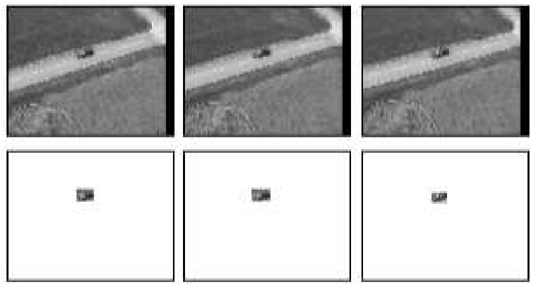


(a) (b)

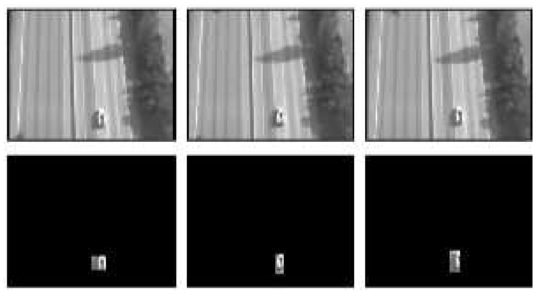
Fig. 4. (a) Corners detection on one aerial image. (b) Description of primitives using corners.

3.2 Experimental results

Fig. 3.2 shows some results with this approach. To verify results of this method, a verifying algorithm based on Dempster-Shafer theory is used.



(a)



(b)

Fig. 5. Example of detection results using "common fate" principle.

References

1. K.L. Boyer and S. Sarkar. *Perceptual Organization for Artificial Vision Systems*. Kluwer, March 2000.
2. P. Burlina, V. Parameswaran, and R. Chellappa. Sensitivity analysis and learning strategies for context-based vehicle detection algorithms. In *DARPA97*, pages 577–584, 1997.
3. J. Canny. A computational approach to edge detection. *PAMI*, 8(6):679–698, November 1986.
4. Schmid Cordelia. Appariement d'images par invariants locaux de niveaux de gris. *PhD thesis, Institut National Polytechnique de Grenoble*, July 1996.
5. C. Harris and M. Stephens. A conbined corner and edge detector. In *Alvey Vision Conference*, pages 147–151, 1988.
6. K. Kaaniche and P. Vasseur. Segmentation and partition of aerial scenes for road traffic diagnosis. In *Asian Conference on Computer Vision*, pages Vol II : 967–972, January 2004.

7. K. Kaaniche and P. Vasseur. Segmentation et partition de squences ariennes pour le diagnostic du trafic routier et autoroutier. In *RFIA, Toulouse*, January 2004.
8. G. Liu and R.M. Haralick. Vehicle ground-truth database for the vertical-view ft. hood imagery. In *ICPR00*, pages Vol I: 342–345, 2000.
9. J. Malik and S. Russell. *A Machine Vision Based Surveillance System for California Roads*. University of California. California Path Research, March 1995.
10. G. Medioni and I. Cohen. Event detection and analysis from video streams. *PAMI*, 23(8):873–889, August 2001.
11. H. Moon, R. Chellappa, and A. Rosenfeld. Performance analysis of a simple vehicle detection algorithm. *IVC*, 20(1):1–13, January 2002.
12. S. Sarkar and P. Soundararajan. Supervised learning of large perceptual organization: Graph spectral partitioning and learning automata. *PAMI*, 22(5):504–525, May 2000.
13. J. Shi and J. Malik. Normalized cuts and image segmentation. *PAMI*, 22(8):888–905, August 2000.
14. P. Soundararajan and S. Sarkar. Investigation of measures for grouping by graph partitioning. In *CVPR01*, pages I:239–246, 2001.
15. Z. Wu and R. Leahy. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *PAMI*, 15(11):1101–1113, November 1993.
16. T. Zhao and R. Nevatia. Car detection in low resolution aerial images. *IVC*, 21(8):693–703, August 2003.