

## ANALYSE DE SEGMENTS TEMPORELS 1-D POUR L'ANALYSE DU TRAFIC ROUTIER

M. Brulin and H. Nicolas

Labri, 351 cours de la Libération, 33405 Talence

University of Bordeaux 1, France

[nicolas@labri.fr](mailto:nicolas@labri.fr)

### Résumé

*La vidéosurveillance est un outil important pour la gestion et la sécurisation du trafic routier. La vidéo surveillance peut être utilisée pour des applications variées telle que le comptage et la classification de véhicules, le calcul de leurs vitesses et comportements. Dans ce contexte, il est souvent important d'être capable d'analyser la vidéo en temps-réel. Néanmoins, en raison d'un nombre croissant de caméras, cet objectif devient plus difficile à atteindre. Une solution consiste à réduire le volume de données à traiter en sélectionnant dans la vidéo les régions d'intérêt, ici essentiellement les zones de routes. Dans cet article, nous proposons d'extraire des segments 1-D d'images dans ces zones d'intérêt et de les analyser pour le comptage et la caractérisation du comportement des véhicules. Les expérimentations réalisées sur des données réelles montrent qu'il est possible d'obtenir, à un faible coût opératoire, des résultats précis.*

### Mots clefs

Vidéosurveillance, trafic routier, segments temporels.

## 1 Introduction

Depuis quelques années, la vidéosurveillance dédiée au trafic routier se développe rapidement et est devenue incontournable pour des questions de sécurité et pour l'analyse et la gestion du trafic [1][2][3]. L'objectif consiste en l'extraction, à partir des données vidéo, d'informations relatives au comportement des véhicules et à l'évolution de la densité du trafic. Pour être réellement efficace, un tel système de vidéosurveillance doit être totalement automatisé et capable de fournir des informations en temps-réel sur le comportement des véhicules. Les événements d'intérêt sont essentiellement : les véhicules entrant ou sortant de la scène, les collisions ou accidents, les véhicules à l'arrêt, la classification des véhicules (voitures, camions, motos ...), la présence de véhicules dans des zones interdites, la détection de trajectoires anormales (par exemple un contre sens), ou

des informations de natures statistiques (estimation de la densité de véhicules, leur vitesse, les changements de voies ...) [4][5][6]. Cela nécessite d'obtenir des informations concernant les contours et textures des véhicules [7], du mouvement [8], ainsi que de leurs trajectoires et leurs vitesses [9].

Pour atteindre la contrainte de temps-réel tout en conservant une analyse efficace, le coût opératoire doit être minimisé. Dans ce contexte, deux solutions peuvent être envisagées. Une première solution consiste à réduire autant que possible les coûts opératoires des phases complexes d'analyse du mouvement et de la segmentation basée objet, mais avec le risque de dégrader trop fortement la qualité des résultats obtenus. Une deuxième solution réside dans la réduction du volume de données devant être analysées. Cette approche est prometteuse dans le contexte de la surveillance du trafic routier pour lequel l'information utile est très localisée dans l'image.

Dans ce contexte, une première solution consiste à définir les régions d'intérêt (ROI) dans les images. Dans le cas de l'analyse du trafic routier, ces zones correspondent aux zones de routes. Les autres régions sont de fait inutiles et peuvent être éliminées. Le sous-échantillonnage temporel est une solution possible pour réduire le volume de données, mais l'accroissement de la distance temporelle qui en résulte rend plus complexe les phases d'analyse spatio-temporelles. Si ces solutions sont intéressantes, elles ne sont pas suffisantes en elle-même. Pour mieux atteindre l'objectif de réduire drastiquement le coût opératoire, il est nécessaire de prendre directement en compte le type d'information réellement utile pour obtenir les informations recherchées. Une solution prometteuse consiste à extraire des vidéos l'évolution temporelle de segments 1-D sélectionnés dans le plan image [10][11]. De tels segments 1-D représentent un volume de données à analyser très faible. S'ils sont choisis judicieusement ils peuvent contenir suffisamment d'information concernant l'analyse du comportement dynamique des véhicules. Dans cet article, nous proposons une méthode permettant de choisir efficacement ces segments et de les analyser pour en extraire des descripteurs utiles pour l'analyse du trafic et le comptage de véhicules.

## 2 Propriétés des segments 1-D pour l'analyse du trafic routier

Dans le cadre d'une application de vidéosurveillance du trafic routier, les régions d'intérêt (ROI) sont les zones de route. Dans la plupart des cas, ces régions sont structurées par des voies de circulation sur lesquelles se déplacent les véhicules. L'évolution temporelle de segments inclus dans ces ROI contient alors des informations utiles. Pour chaque segment, une image, appelé ici IST (Image d'un Segment Temporel) est construite en accumulant les segments au fil du temps. La figure 1 montre deux exemples d'IST obtenus avec des segments parallèles et perpendiculaires aux voies de circulation. Dans la suite de cet article, pour une meilleure visualisation des IST, l'effet de perspective est compensé en utilisant un rectangle (dans l'espace 3-D) défini par les voies de circulation parallèles. Les Figures 2 et 5 montrent des exemples d'une telle compensation. Les caractéristiques principales d'une IST issues d'un segment parallèle aux voies de circulation sont les suivantes :

Il peut être observé que chaque véhicule en mouvement est représenté dans une IST par une bande démarant généralement du bas à gauche vers le haut à droite dans l'image (voir la figure 1, bas). Le nombre de bandes caractérise le nombre de véhicules passant sur le segment durant la période temporelle considérée. La vitesse et l'accélération de chaque véhicule est obtenue en calculant les dérivées premières et secondes du contour de chaque bande. Si la vitesse est constante, la bande est une ligne droite et son orientation donne la vitesse du véhicule. A un instant  $t$  donné, la largeur de la bande représente la longueur ou la hauteur (selon l'orientation de la caméra) du véhicule. La bande peut ne pas être présente du bas vers le haut de l'image en cas de changement de voie du véhicule. La couleur dominante de la bande correspond généralement à la couleur dominante du véhicule.

Une telle IST contient donc a priori les informations suffisantes pour une analyse du comportement dynamique des véhicules.

## 3 Méthode proposée

La figure 3 propose le schéma général de la méthode proposée. Elle est décomposée en deux phases principales. La première consiste en une phase de pré-analyse consistant en une estimation de l'arrière-plan de la scène, de la détection des voies de circulation, et de la sélection des segments utilisés pour construire les IST. La seconde consiste en une phase d'analyse temps-réel pour l'analyse dynamique des véhicules et leur comptage. Deux hypothèses sont introduites : la caméra est supposée être fixe et approximativement orientée le long de l'axe de la route.

Elles sont souvent raisonnablement remplies dans le cas de caméras de surveillance du trafic autoroutier.

### 3.1 Phase de pré-analyse

L'arrière-plan de la scène est obtenu en utilisant une approche basée pixel de mélange de gaussiennes [12]. On modélise l'historique des valeurs d'un pixel par plusieurs distributions ce qui permet d'être robuste par rapport aux changements d'illumination et aux passages de véhicules en mouvement. Les paramètres (poids  $w$ , moyenne  $\mu$  et covariance  $\sigma$ ) sont mises à jour dynamiquement. La probabilité  $P$  d'occurrence d'une couleur  $u$  au pixel  $p$  à l'instant  $t$  est donné par (avec  $k$  le nombre de gaussiennes):

$$P(I_{p,t} = u) = \sum_{i=1}^k (w_{i,p,t} \mathcal{N}(I_{p,t}, \mu_{i,p,t}, \sigma_{i,p,t}))$$

La gaussienne la plus vraisemblable est sélectionnée si  $d < 2.5\sigma$ . Les paramètres de la gaussienne sélectionnée  $k$  sont mises à jour avec :

$$\begin{aligned} w_{k,p,t} &= (1 - \alpha)w_{k,p,t-1} + \alpha \\ \mu_{k,p,t} &= (1 - \rho)\mu_{k,p,t-1} + \rho I_{p,t} \\ \sigma_{k,p,t}^2 &= (1 - \rho)\sigma_{k,p,t-1}^2 + \rho(I_{p,t} - \mu_{k,p,t})(I_{p,t} - \mu_{k,p,t})^T \end{aligned}$$

$\alpha$  et  $\rho$  sont des coefficients constants. Pour les gaussiennes non sélectionnées, seul le poids est mise à jour de la manière suivante :

$$w_{i,p,t} = (1 - \alpha)w_{i,p,t-1} - \alpha(1 - c)$$

$c$  est empiriquement fixé à 0.1. Si une gaussienne n'est pas sélectionnée durant une certaine période de temps, son poids devient négatif, elle est alors supprimée. Il n'est donc pas nécessaire de fixer un nombre de gaussiennes maximum.  $\rho$  est mise à jour plus rapidement pour les gaussiennes récemment créés car elle sont moins stables que les gaussiennes pour lesquelles de nombreuses estimations sont disponibles. De ce fait nous avons :

$$\rho = \alpha + (1 - \alpha)/n.$$

Avec  $n$  le nombre de valeurs utilisées pour construire la gaussienne. Finalement, pour chaque pixel, l'arrière-plan est calculé en utilisant la gaussienne ayant le plus fort ratio  $w/\sigma$ . Un pixel pour lequel une autre gaussienne a été sélectionné est considéré comme ne faisant pas partie de l'arrière-plan. Un filtrage morphologique est utilisé pour remplir les trous et pour éliminer les toutes petites régions.

Les voies de circulation sont ensuite détectées en utilisant une méthode basé sur l'algorithme CHEVP. Pour cela, un détecteur de contours de canny est appliqué. Ensuite, les paramètres de droites sont estimés en utilisant la

transformée de Hough. Le point de fuite des lignes séparant les voies est alors estimés, en utilisant les intersections  $\gamma$  des lignes précédemment estimées, en utilisant le critère suivant :

$$VP = \operatorname{arg\,min}_{p \in J} \sum_{\gamma \in I} \operatorname{DIST}(p, \gamma)$$

Où  $I$  est l'ensemble des intersections.  $J$  est le plus petit cercle dans le plan image incluant l'ensemble des intersections. Les lignes n'intersectant pas ce cercle  $C$  centré sur le point de fuite sont éliminés. Le diamètre du cercle a été empiriquement fixé à 10% de la largeur de l'image. Nous obtenons ainsi une segmentation des zones d'intérêt (voir la Figure 4). Les résultats expérimentaux montrent que la plupart des lignes séparant les voies de circulation sont correctement estimées. Un segment est alors automatiquement sélectionné sur le milieu de chaque voie et dans la zone la plus proche de la camera (partie basse de l'image) ainsi détectée (voir la figure 4)

### 3.2 Analyse et Classification des IST

L'objectif consiste tout d'abord à séparer l'avant-plan de l'arrière-plan dans chaque image IST. Chaque pixel devant être classifié comme *Background* ( $B$ ) ou *Foreground* ( $F$ ). En pratique, à chaque instant un nouveau segment  $S_t$  est ajouté à l'image IST. Chaque pixel dans  $S_t$  doit alors être classifié en fonction de la segmentation obtenue pour les segments précédents. Cela est réalisé en utilisant l'algorithme suivant :

- *Prédiction temporelle* : Par construction, les pixels situés sur la même ligne horizontale d'une IST représente le même point physique au cours du temps (la caméra est supposée fixe). Le segment  $S_{t-1}$  à l'instant  $t-1$  est alors projeté à l'instant  $t$  en supposant que l'accélération est nulle et la vitesse estimé de l'objet (à l'instant  $t-1$ , voir ci-dessous). Le reste du segment à l'instant  $t$  est prédit comme faisant partie de l'arrière-plan.

- *Segmentation spatiale* : La méthode de segmentation basée sur le mélange de gaussienne présentée ci-dessus est réutilisée pour obtenir une classification basée pixel du segment  $S_t$  (avec les deux labels  $F$  et  $B$ ). Un opérateur morphologique 1-D est alors utilisé pour éliminer les pixels  $F$  ou  $B$  isolés. Chaque sous-segment défini par une fenêtre glissante  $w$  est alors classifié comme *Foreground* si :

$$\operatorname{Card}\{p \in w / s(p) = F\} > \operatorname{Card}\{p \in w / s(p) = B\}$$

Où  $S(p)$  est le label du pixel  $p$ . La longueur de  $w$  est définie par la moyenne constatée expérimentalement pour les véhicules du segment considéré. Cela permet d'obtenir des sous-segments classifiés comme *Foreground*. Les pixels aux frontières du sous segment sont récursivement éliminés s'ils étaient classifiés comme *Background* par la Segmentation spatiale.

- *Classification finale* : Elle est obtenue après application des critères suivants :

- 1- Sous-segments en recouvrement (partiel ou total) classifiés comme *foreground* soit par la segmentation spatiale, soit par la prédiction temporelle : Leur union est classifiée comme *Foreground*. Cela permet un suivi temporel du véhicule.
- 2- Sous-segment classifié comme *foreground* par la prédiction temporelle qui n'intersecte aucun segment classifié comme *Foreground* par la segmentation spatiale : Il n'est pas validé comme segment *Foreground* à l'instant  $t$ . Cela signifie généralement que le véhicule a changé de voie de circulation.
- 3- Sous-segment classifié comme *Foreground* par la segmentation spatiale ne recouvrant aucun sous-segment *foreground* issue de la prédiction : ce segment est considéré comme un nouveau véhicule.

### 3.3 Applications des IST à l'analyse du trafic

Différents descripteurs peuvent être obtenus à partir des images IST:

- *Estimation de la vitesse des véhicules*: Pour chaque bande détectée dans une IST, la vitesse du véhicule est théoriquement donnée par la dérivée spatiale du contour de la bande. Si la vitesse est constante, ce contour est une ligne droite. Pour éviter les problèmes de perspective, le contour utilisé est celui correspondant à l'arrière du véhicule si celui-ci s'éloigne de la caméra (ou à l'avant du véhicule dans le cas inverse).

- *Détection d'une congestion de trafic*: Une congestion du trafic, préfigurant un bouchon est défini lorsque la vitesse passe en-dessous d'un seuil de vitesse bien définie.

- *Détection de véhicules à l'arrêt*: Si une bande devient horizontale, cela signifie que le véhicule est à l'arrêt (en pratique un seuil de vitesse très faible peut être choisi). Si des véhicules sont simultanément détectés à l'arrêt dans toutes les voies, une situation de bouchon est ainsi détectée.

- *Comptage de véhicules*: Le nombre de véhicules est directement fourni par le nombre de bandes. A noter que lorsqu'un bouchon est détecté, les véhicules étant très proches les uns des autres, le comptage devient inopérant car les bandes ne peuvent plus être clairement séparées.

- *Détection des dépassements et changements de voies*: Les cas 2 et 3 décrit dans la section 3.2 correspondent théoriquement à l'apparition ou la disparition d'un véhicule dans une voie donnée. Cela correspond de fait à des bandes incomplètes dans l'IST.

#### 4. RESULTATS EXPERIMENTAUX

La méthode proposée a été testée sur un corpus de 25 vidéos de surveillance (avec des durées de 10 à 30 minutes) obtenues sur des autoroutes variées avec des conditions variable d'illumination et de conditions d'éclairage. Après la phase de pré-analyse, toutes les phases d'analyse sont calculées en temps réel. Les résultats suivants ont été obtenus:

- *Détection d'une congestion de trafic*: Ce type de situation est bien détecté dans tous les cas disponible dans les vidéos de test. Voir l'exemple de la figure 5. Lorsque les véhicules se déplacent très lentement (ou sont à l'arrêt), leurs distances relatives sont réduites, leurs bandes respectives peuvent alors facilement être fusionnées ce qui rend impossible le comptage.

- *Détection de véhicules à l'arrêt*: Les véhicules à l'arrêt sont systématiquement détectés (voir les figures 5 et 6).

- *Comptage de véhicules*: un rappel de 94% et une précision de 87 % sont obtenus (résultats moyens pour chacune des voies de circulation). Les principaux problèmes sont liés aux situations dans lesquelles des camions recouvrent plusieurs voies de circulation dans le plan image. En conséquence, ils peuvent être comptés deux fois.

- *Détection des changements de voies*: Un rappel de 97% et une précision de 89 % sont obtenus.

L'utilisation de ces descripteurs de base permet d'étudier le comportement des véhicules d'une manière plus globale. Cela est possible en créant des scénarios à partir d'une succession d'événement fondamentaux. La figure 6 donne un exemple typique: Dans cette vidéo, un véhicule stoppe au milieu de la route. Un autre camion arrive et le double puis revient sur sa voie initiale. Finalement le véhicule à l'arrêt redémarre. Tous ces événements ont été correctement et automatiquement détectés dans l'ordre chronologique correct:

- Entrée dans la scène du premier camion (Fig. 5a).
- Entrée dans la scène du second véhicule (une voiture, Fig. 5b).
- La voiture est détectée à l'arrêt (Fig. 5c).
- Entrée d'un troisième véhicule.
- Disparition de ce véhicule de sa voie de circulation avant d'arrivée près du véhicule à l'arrêt (Fig. 5d).

- Apparition d'un véhicule dans la voie de circulation voisine.
- Disparition de ce véhicule de cette seconde voie de circulation.
- Apparition de ce véhicule dans la première voie au-delà du véhicule à l'arrêt (Fig. 5e)
- Redémarrage du véhicule à l'arrêt (Fig 5f).

#### 5. CONCLUSION ET PERSPECTIVES

Cet article propose une nouvelle méthode permettant le comptage et l'analyse du comportement de véhicules dans le cadre de la surveillance et l'analyse temps-réel du trafic routier. Elle est basée sur l'analyse temporelle de segments 1-D situés dans chacune des voies de circulation. Elle permet de compter les véhicules, de détecter les situations de congestion de trafic, de détecter les véhicules à l'arrêt et les bouchons, ainsi que les changements de voies. L'obtention de ces descripteurs de base peut alors permettre de caractériser des comportements plus complexes de véhicules. L'analyse se fait en temps-réel après la phase initiale. L'approche a été validée sur des données vidéo réelles obtenues dans le cadre de vidéo surveillance du trafic routier.

Plusieurs perspectives de ce travail sont possibles tels que: la détection de petits véhicules tels que les motos, vélos ou piétons, une meilleure gestion des véhicules projetés sur plusieurs voies de circulation dans le plan image (due à l'effet de perspective), et la définition exhaustive de scénario de comportements complexes utiles pour la gestion et la sécurisation du trafic.

#### Références

- [1] V. Kastrina, M. Zervakis and K. Kalaitzakis. *A survey of video processing techniques for traffic applications*. Image and Vision Computing, Vol. 21, N°4, pp. 359-381, 2003.
- [2] N. Buch, S.A. Velastin and J. Orwell. *A review of computer vision techniques for the analysis of urban traffic*. IEEE Transactions on Intelligent Transportation system, Vol. 12, N°3, pp. 920-939, 2011.
- [3] B. Tian, Q. Yao, Y. Gu, K. Wang and Y. Li. *Video processing techniques for traffic flow monitoring: A survey*. In Proc. of Int. Conf. on Intelligent Transportation Systems, pp. 1103-1108, 2011.
- [4] Z. Zhu, G. Xu, B. Yang, D. Shi and X. Lin. *VISATRAM: A real-time vision system for automatic traffic monitoring*. Image and Vision Computing, Vol. 18, No. 10, pp.781-794, 2000.

[5] A. Yoneyama, C.H. Yeh and C.C.J. Kuo. *Robust vehicle and traffic information extraction for highway surveillance*. Image and Vision Computing, Vol. 2005, pp. 2305-2321, 2005.

[6] T. Rodriguez and N. Garcia. *An adaptive, real-time, traffic monitoring system*. Machine Vision and Applications. Vol. 21, No. 4, pp. 555-576, 2010.

[7] A. Bissacco, P. Saisan and S. Soatto. *Gait recognition using dynamic affine invariant*. In int. Symposium on Mathematical Theory of Network, and Systems. 2004.

[8] A. Adam, E. rivlin, I. Shimshoni and D. Reinitz. *Robust real-time unusual event detection using multiple fixed-location monitors*. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 30, N°3, pp.555-560, 2008.

[9] C. Stauffer. *Estimating tracking sources and sinks*. In Proc. of Computer Vision and Pattern Recognition, IEEE, Vol. 4, pp.35-45, 2003.

[10] Y. Malinovski, Y. Wang and Y.J. Wu. *Video-based vehicle detection and tracking using spatio-temporal maps*. Proc. of the Annual Transportation Research Board meeting, Washington DC, 2009.

[11] Z. Zhu, G. Xu, B. Yang, D. Shi and X. Lin. *VISITRAM: A real time vision system for automatic traffic monitoring*.

[12] Bouwmans, T., El Baf, F. and Vachon, B., *Background modeling using mixture of gaussians for foreground detection - A survey*. Recent Patents on Computer Science, pp. 219-237, 2008.

Fig. 1: Haut: segment 1-D perpendiculaire aux voies de circulation. Bas: Segment 1-D parallèle aux voies de circulation. Axe horizontal axis: axe temporel. Axe vertical: Axe du segment.



Fig. 2: IST avant (haut) et après (bas) la Compensation de l'effet de perspective

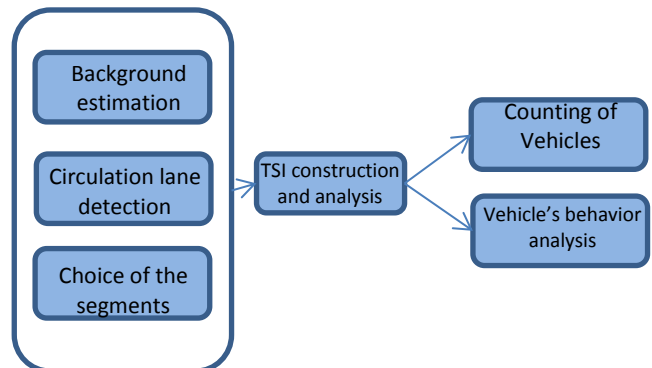
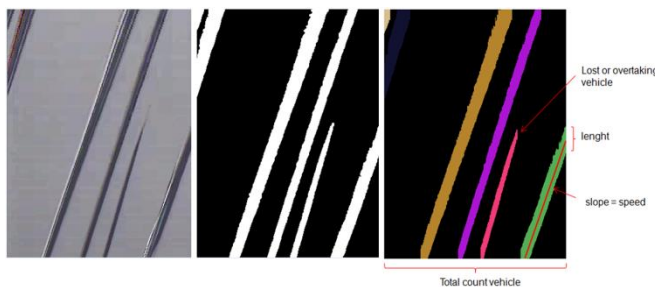
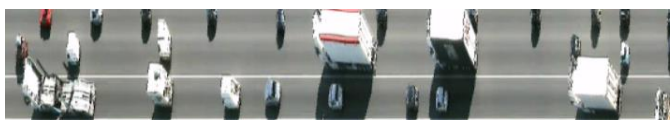


Fig. 3: Schéma général de la méthode proposée



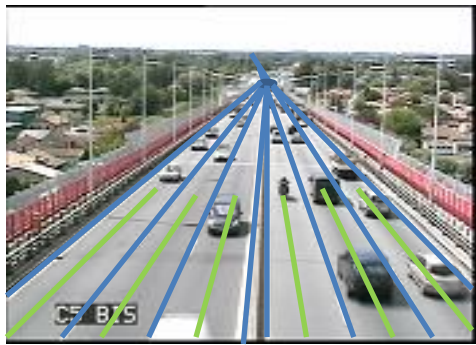


Fig. 4: Voies de circulation et point de fuite estimés (bleu), segments (vert)

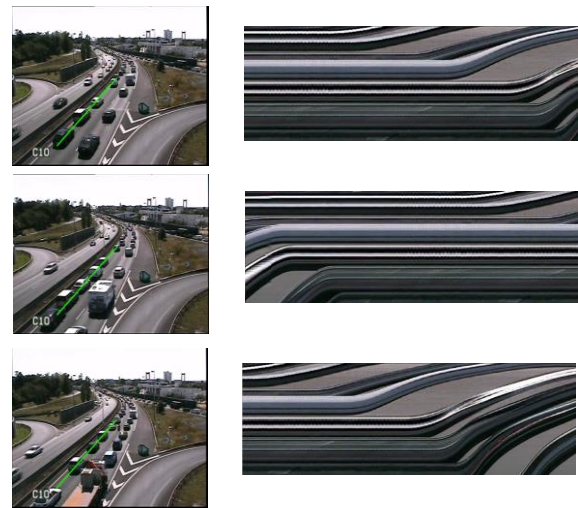


Fig. 5: Exemple pour lequel une congestion de trafic a été automatiquement détectée

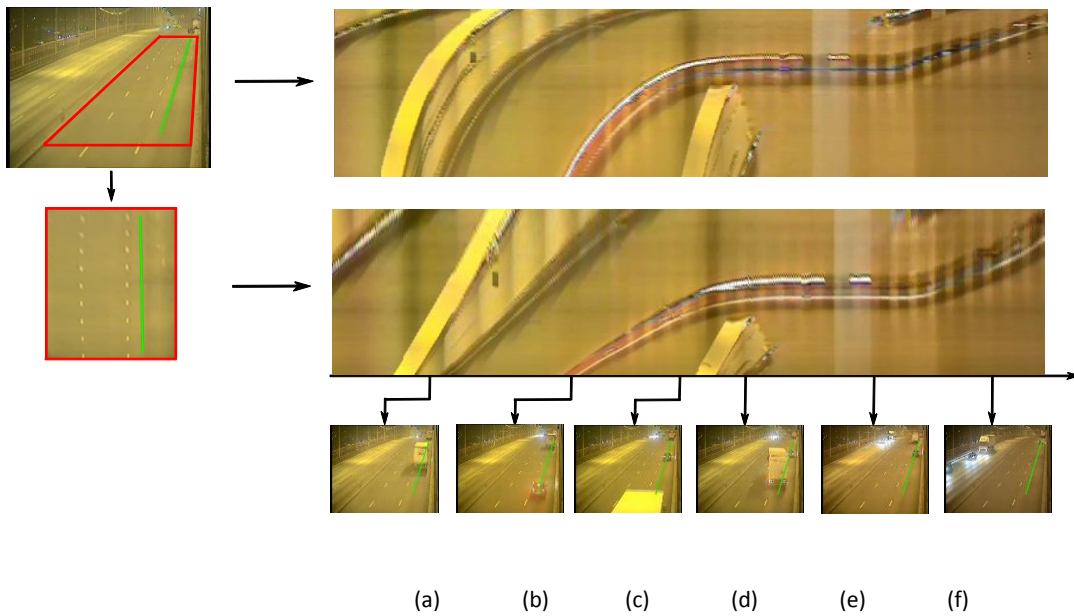


Fig. 6: Haut: original image et l'IST correspondant au segment en vert. Le rectangle rouge est utilisé pour la compensation de l'effet perspective. Milieu : IST après compensation de la perspective. Bas: images originales illustrant les évènements successifs