

# Poursuite de Véhicules en Scènes Autoroutières

J. BECERRA<sup>1</sup> D. IZQUIERDO<sup>2</sup> Y. BERTHOUMIEU<sup>2</sup> M. DONIAS<sup>1</sup>

<sup>1</sup> UMR 5131 ESI/LAP – ENSEIRB    <sup>2</sup> UMR 5818 IXL – ENSEIRB  
1, Avenue du docteur Albert Schweitzer, BP 99, 33402 Talence cedex – France

{javier.becerra, david.izquierdo, Yannick.Berthoumieu, marc.donias}@enseirb.fr

**Résumé** – Dans cet article, nous proposons une nouvelle méthode de suivi de véhicules dédiée à la surveillance autoroutière. Le processus de suivi s’appuie sur une segmentation spatio-temporelle caractérisée par la coopération de deux techniques distinctes. La première d’entre elles est une technique de bas niveau qui a pour rôle de séparer l’ensemble des pixels associés aux objets en mouvement de ceux statiques. A cet effet, une nouvelle méthode de génération d’une image de référence adaptative qui prend en compte les changements d’éclairage ambiant a été développée. Les perturbations induites par les phénomènes d’ombrage sont atténuées par l’utilisation d’un modèle colorimétrique. La seconde technique employée consiste en une méthode discriminante en terme de modèle de mouvement. Il s’agit d’une méthode de haut niveau qui s’appuie sur un ensemble de descripteurs associés à chaque objet. Utilisant un algorithme de type Expectation-Maximisation (EM), cette approche permet également de gérer les problèmes d’occlusion. Concernant la poursuite, l’identification de chaque objet est ensuite assurée par une étape de reconnaissance dans l’espace de classification engendré par les descripteurs. L’efficacité des procédés décrits est démontrée dans un contexte réel sur quelques exemples de scènes autoroutières.

**Abstract** – We introduce a new vehicle tracking algorithm focused on traffic surveillance. The tracking process is based on a spatiotemporal segmentation issued from the collaboration of two different techniques. The first one is a low level treatment that classifies pixels as belonging to moving objects or background. This is accomplished via a new method used to generate an adaptive reference image and that is able to deal with ambient light variations. Also at this level, a colorimetric model is used to minimise the perturbations induced by shadow phenomena. The second technique is a high level one, using a set of descriptors associated to each object and an Expectation-Maximisation (EM) algorithm based on movement. EM assures the robustness of the system regarding occlusions. At the pursuit stage, the identification of the objects takes place in the feature space of the descriptors. We show the efficiency of these methods in a real-world context using several examples of traffic scenes.

## 1. Introduction

Durant la dernière décennie, de nombreuses méthodes de segmentation basée sur le mouvement ont été proposées dans la littérature. Certaines d’entre-elles qui s’intéressent spécifiquement au domaine de la surveillance autoroutière peuvent être regroupées en plusieurs grandes familles. Une première famille concerne les techniques de segmentation utilisant des primitives associées au Mouvement Apparent (MA) [1,2,3]. Ces méthodes s’appuient sur des hypothèses statistiques permettant de modéliser le contexte d’analyse par un processus de Markov ou de l’exprimer à travers une fonction de vraisemblance. La modélisation utilisée dans cette famille exploite un mélange statistique traité généralement par une approche de type Expectation-Maximization (EM) [3]. Une deuxième famille se base sur la détection des changements existant entre l’image courante et une Image de Référence Adaptative (IRA) [4,5]. En générant une image de référence qui ne contient que la partie statique de la scène ou « information du fond », ce principe conduit à une extraction rapide des objets en mouvement. Dans ce contexte d’approches MA et IRA, nous proposons une méthode complète de segmentation utilisant leurs avantages respectifs. Notre approche utilise une représentation de haut niveau s’appuyant sur plusieurs descripteurs représentant chaque objet. Cet ensemble de descripteurs sera dans la suite

de l’article désigné par le terme modèle/objet. L’identification des modèles/objets est établie à partir des règles de mise en correspondance associées à leur espace de représentation. L’actualisation de cet espace de classification dépend des résultats du processus de segmentation utilisant les approches IRA et MA. Dans ce sens, dans un premier temps, et dans l’optique de renforcer les performances du système, nous avons introduit une nouvelle méthode IRA moins sensible aux distorsions photométriques dues aux conditions atmosphériques naturelles. Dans un second temps, un modèle colorimétrique de l’ombre associée à l’objet est établi [6]. Ce modèle nous permet de stabiliser la segmentation de la scène au sens de l’objet par suppression de l’ombre qui est toujours incluse dans la segmentation donnée par la méthode IRA.

Cet article est articulé de la manière suivante. La section 2 décrit un nouvel algorithme de segmentation temporelle. La section 3 explicite le processus de suppression d’ombre utilisé. Dans la section 4, nous décrivons le processus de suivi qui prend en compte les Attributs d’Objet et qui est supervisé par un algorithme EM. La section 5 présente quelques exemples expérimentaux obtenus sur des scènes autoroutières réelles. La conclusion clôture cet article en section 6.

## 2. Segmentation temporelle fondée sur une nouvelle approche IRA

La détection de variations et la segmentation fondée sur le mouvement sont des tâches fondamentales pour tout système de vidéo surveillance automatique. Avant l'étape de suivi, une réduction de l'observation, ou information utile, peut être obtenu sous la forme de masques binaires associés aux régions connexes issus de la segmentation temporelle. A cette fin, la bibliographie en vision par ordinateur offre une grande variété de techniques [5,7]. L'une d'entre-elles nommée Image de Référence Adaptative (IRA) se dégage tout particulièrement [4]. Dans la littérature, cette technique se fonde sur la formulation suivante:

$$B^k = \alpha^{k-1} B^{k-1} + (1 - \alpha^{k-1}) I^{k-1}, \quad (1)$$

où  $B$  représente l'image de référence,  $I$  est l'image courante,  $\alpha$  représente la mémoire du système (la capacité du système à répondre aux changements) et  $k$  est l'index de l'image traitée. Cette approche est adaptée aux applications à flux continu. Elle est en effet caractérisée par un faible coût de calcul. Par contre, ce type de mise à jour est inefficace lors de changements rapides de luminosité ambiante. C'est notamment le cas lors de passages nuageux, au lever ou au coucher du soleil. Les figures 1 et 2 montrent le profil de la courbe de luminosité  $I^k(p)$  pour un pixel  $p$  donné durant le passage d'un objet.

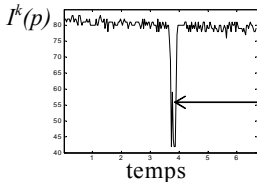


FIG. 1 : luminosité ambiante constant.

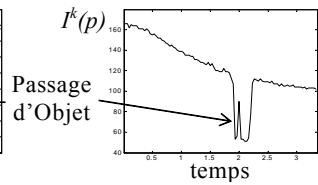


FIG. 2 : changement de luminosité ambiante.

Afin de détecter tous les passages d'objets et de réduire l'influence des phénomènes naturels, nous proposons un nouveau procédé présentant une mise à jour robuste de la référence fondée sur :

- **une procédure de décision pour le passage d'objets;** celle-ci exploite la différence existant entre les caractéristiques de la courbe de luminosité ambiante et celle des objets en utilisant une approche dérivative (section 2.1). Cette approche se fonde sur la forte variation du profil de luminosité causée par le passage d'un objet. La méthode de détection s'appuie sur une formulation de type variationnelle ;
- **l'utilisation de deux références auxiliaires;** les images de référence  $B_1^k$  et  $B_2^k$  sont associées à la pente de la courbe de luminosité  $I^k(p)$  (section 2.2). Elles sont introduites pour s'adapter au mieux à une variation de la luminosité du fond de l'image.

### 2.1 Approche dérivative

L'introduction d'une vision "forward" est impérative pour assurer une bonne robustesse par rapport aux changements de luminosité engendrés par le passage d'un objet. Cette

observation non causale permet de prédire le passage d'un objet par une approche variationnelle.

$$d^k = \sum_{i=0}^n |I^{k+i+1} - I^{k+i}| \quad (2)$$

La décision du passage de l'objet  $D^k(p)$ , obtenue en un pixel de position  $p$ , implique l'utilisation d'un seuil sur l'opérateur de variationnelle,  $d^k(p)$ :

$$D^k(p) = \begin{cases} 1 & \text{si } d^k(p) \leq T^k(p) \\ 0 & \text{ailleurs} \end{cases} \quad (3)$$

$T^k(p)$  est un seuil adaptatif défini comme le produit d'une constante  $c$  et  $m_2^k(p)$  le moment de deuxième ordre associé à  $d^k(p)$ , dont la mise à jour suit l'équation générale:

$$T^k(p) = c \cdot \sqrt{m_2^k(p)}, \quad (4)$$

$$m_2^k(p) = m_2^{k-1}(p)\beta(p) + (1 - \beta(p))(d^{k-1}(p))^2 \quad (5)$$

où  $\beta(p)$  permet une mise à jour sélective définie par  $\alpha$ , la mémoire du système, et  $D^k(p)$ :

$$\beta(p) = 1 - \alpha D^k(p) \quad (6)$$

### 2.2 Mise à jour de la référence

Dans un contexte de scènes extérieures, le système de détection peut-être confronté à une variation progressive de luminosité de l'image. La référence doit donc être mise à jour en prenant en compte ce gradient. Pour assurer la meilleure poursuite nous proposons de tester deux hypothèses de maintien de la référence afin de restituer le niveau correct de l'image du fond. Ce principe permet d'inhiber l'état de détection du mouvement. La mise à jour de ces références n'est activée que lors du passage d'un objet. Un schéma récursif extrait l'image de référence  $B^k$  à l'aide de deux images de référence auxiliaires différentes  $B_1^k$  et  $B_2^k$ . Lors d'un passage d'un objet sur un pixel  $p$ , les valeurs de  $B_1^k$  et  $B_2^k$  deviennent :

$$\begin{aligned} B_1^k &= B_1^{k-1} \\ B_2^k &= B_2^{k-1} + L^{k-1} \end{aligned} \quad (7)$$

où  $L^{k-1}$  représente la valeur estimée de la pente de la courbe  $B_2^k$  :

$$L^k = \rho L^{k-1} + (1 - \rho)(B_2^k - B_2^{k-1}) \quad (8)$$

Au pixel  $p$ , la valeur finale de  $B^k$  retenue est celle qui minimise l'écart à la valeur de luminance :

$$B^k = \text{Arg} \min_{B \in \{B_1, B_2\}} (|B^k - I^k|) \quad (9)$$

Dans le cas particulier d'un pixel ne correspondant pas à un passage d'objet, chacune des références est tout simplement prise égale à la valeur de l'image courante :

$$B^k = B_1^k = B_2^k = I^k \quad (10)$$

## 2.3 Processus IRA

Notre implantation s'appuie sur trois images consécutives (du fait du choix de  $n=1$  dans l'équation du variationnelle (2)). La figure 3 résume le processus IRA au complet. On définit la sortie du processus complet  $Ro^k = \{R_{o_i}^k\}_{i=1, \dots, r}$  comme l'ensemble des  $r$  régions connexes dans la  $k^{\text{ième}}$  image qui sont obtenues à partir du résultat de la différence seuillée entre  $B^k$  et  $I^k$  [8].

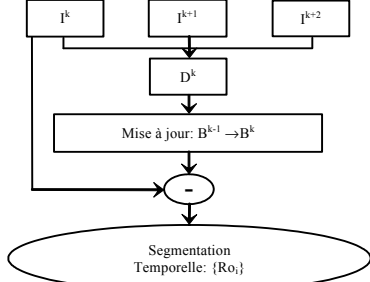


FIG. 3 : Diagramme complet de la Segmentation Temporelle.

## 3. Suppression de l'ombre

Dans le cas d'images en couleur, le passage d'un point éclairé à un point ombré se traduit par une chute très nette de la luminosité alors que les valeurs relatives de chrominance rouge et verte sont conservées [6]. Dans ce contexte, nous avons implanté un test multicritère pour classer un point de l'image  $I^k(p)$  comme étant un point ombré. L'information relative à la luminosité est extraite à partir de la composante  $Y$  de l'espace  $YCbCr$ , et celle relative à la couleur à partir des composantes  $r$  et  $g$  de l'espace  $rgb$  normalisé. Un pixel est étiqueté comme associé à l'ombre dès lors que les trois équations suivantes sont vérifiées :

$$\begin{aligned} \frac{Y_I(p)}{Y_B(p)} &< \alpha_Y, \\ |r_I(p) - r_B(p)| &< \beta_r, \\ |g_I(p) - g_B(p)| &< \beta_g, \end{aligned} \quad (11)$$

où  $\alpha_Y$ ,  $\beta_r$  et  $\beta_g$  sont déterminés empiriquement. Pour régler ses valeurs, nous nous sommes basés sur les différences mesurées dans plusieurs séquences autoroutières.

La suppression de l'ombre sur les différentes régions  $Ro^k$  produit un nouvel ensemble de régions  $R^k$ . Cet ensemble est utilisé comme l'entrée du processus de suivi.

## 4. Processus de suivi

L'objectif du processus de suivi est l'identification de chaque objet en mouvement présent dans la scène. Cet objectif est atteint en plusieurs étapes (cf. fig. 4). Il s'agit de mettre en correspondance un ensemble de descripteurs associé à chacun des objets. La description de haut niveau de chaque objet  $j$  présent dans la  $(k-1)^{\text{ième}}$  image est donnée par un ensemble de modèles notés  $M_j^{k-1}$ . Chaque  $M_j^{k-1}$  comprend un ensemble d'Attributs d'Objets (barycentre, aire, contour et

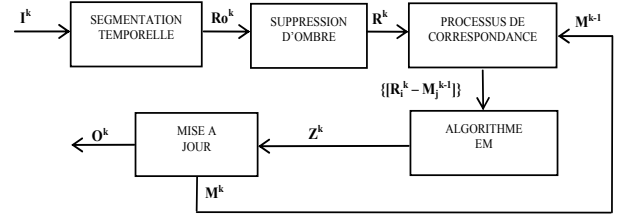


FIG.4 : Diagramme du processus complet de suivi.

modèle de mouvement à 6 paramètres  $\theta_j^{k-1}$ ) permettant l'identification de l'objet dans l'image suivante à partir de l'ensemble des régions  $R_i^k$ . L'identification entre régions et modèles est réalisée par un *Processus de Correspondance*. Il s'agit de comparer chaque région  $R_i^k$  avec la prédiction du modèle  $\tilde{M}_j^k$ . Ce modèle prédit est extrait à partir du modèle  $M_j^{k-1}$  au moyen d'une compensation de mouvement,  $\tilde{M}_j^k = [M_j^{k-1}]_{o_{k-j}}$ , prenant en compte les paramètres  $\theta_j^{k-1}$  estimés dans l'image  $(k-1)^{\text{ième}}$  [1,9].

Lors de la phase de mise en correspondance, les régions obtenues par la segmentation temporelle peuvent n'être que des portions des objets attendus (figure 5) ou correspondre à deux objets différents (figure 6). La phase de reconstruction s'appuyant sur une approche EM décrite dans la section 4.1 se révèle donc nécessaire pour identifier les zones de l'image correspondant à chaque objet.

### 4.1 Algorithme EM

L'algorithme Expectation-Maximisation (EM) met en œuvre une segmentation de l'image fondée sur une procédure de mélange de classes [3]. Dans cette procédure, chaque classe du mélange est associée à un modèle de mouvement à six paramètres [1,9]. De cette façon, nous calculons une probabilité (à posteriori) d'appartenance aux classes pour chaque pixel. Cette probabilité est modélisée par une distribution gaussienne définie par les modèles de mouvement. La segmentation spatiale est obtenue par association des pixels qui ont une probabilité maximale par rapport au modèle de mouvement.

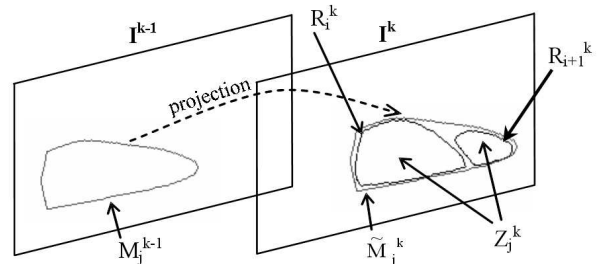


FIG. 5 : Modèle:  $M_j^{k-1}$ , Modèle projeté:  $\tilde{M}_j^k$ , Régions:  $R_i^k, R_{i+1}^k$  et la Zone,  $Z_j^k$  qui correspond à l'objet réel.

L'algorithme extrait une zone  $Z_j^k$  regroupant les régions  $R_i^k$  (ou extrait des régions en cas d'occultation) qui correspondent avec le modèle prédit  $\tilde{M}_j^k$ . Chacune de ces zones identifie la localisation d'un objet sur l'image (fig. 5).

Les attributs  $M_j^k$  de chaque modèle sont mis à jour après cette étape du processus EM. Pour estimer les paramètres du modèle de mouvement, nous utilisons une méthode LMS incrémentale [1,9].

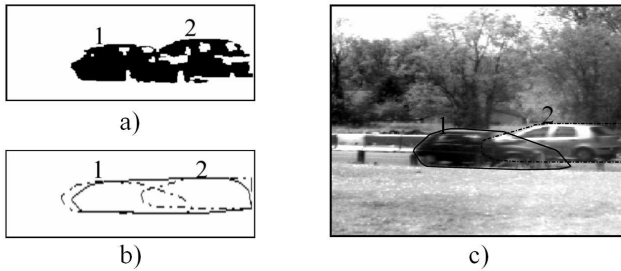


FIG. 6 : Image réelle avec traitement d'occlusion.



FIG.7 : Suivi avant (a) et après (b) suppression de l'ombre.



FIG. 8 : Suivi d'objet dès son entrée dans le champ de vision.

## 5. Résultats

De façon à tester l'algorithme dans des conditions de variations de luminosité, nous avons choisi une séquence réelle d'un milieu autoroutier comportant des passages nuageux et une seconde comportant des objets ombrés.

La figure 6a présente un cas d'occlusion entre deux objets : la segmentation primaire représentée, issue de la méthode IRA, ne lève pas l'ambiguïté car les objets se recouvrent, formant une seule région. La figure 6b contient une unique région  $R_i^k$  et deux modèles  $M_j^{k-1}$  qui la séparent en deux parties distinctes. Finalement, la figure 6c montre l'image réelle avec les deux modèles  $M_j^k$  actualisés. Par ailleurs, la figure 7 présente des résultats de suivi obtenus avant et après suppression d'ombre à l'aide du modèle colorimétrique (11). Le suivi d'un objet dès son entrée dans l'image peut être visualisé dans la figure 8.

## 6. Conclusion

Dans cet article, nous avons proposé un processus complet pour le suivi automatique de véhicules en scènes autoroutières. Les atouts de notre schéma sont sa robustesse aux changements des conditions de luminosité ambiante, le processus de suppression d'ombre sur les objets ainsi que sa capacité de gérer les cas d'occlusion, sur – et – sous – segmentations. Dans un premier temps une procédure IRA non supervisée extrait les régions en mouvement dans l'image. Le processus de suppression d'ombre permet une meilleure localisation de l'objet et évite les ennuis provoqués par la détection de l'ombre faisant partie de l'objet. Dans un second temps une procédure de discrimination est mise en œuvre mettant en compétition cette segmentation avec un ensemble de descripteurs associés à chacun des objets. Ce schéma conduit à une méthode performante de suivi grâce à la modélisation géométrique et à

l'identification fondée sur le mouvement propre à chaque véhicule. La mise en correspondance est assurée par une procédure de type EM.

## 7. Références

- [1] J.M. Odobez and P. Bouthemy, *Direct incremental model-based image motion segmentation analysis for video analysis*, Signal Processing, vol.66, pp. 143-155, 1998.
- [2] C. Dumontier, F. Luthon, J.P. Charras, *Real Time DSP Implementation for MRF-Based Video Motion Detection*, IEEE Transactions on Image Processing, vol.8 n°10, pp.1341-1347, October 1999.
- [3] H. Sawhney and S. Ayer, *Compact representation of videos through dominant and multiple motion estimation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(8), August 1996.
- [4] D. Koller, J. Weber and J. Malik, *Robust Multiple Car Tracking with Occlusion Reasoning*, 3rd ECCV '94, Stockholm Sweden, pp. 189-196, May 1994.
- [5] C. Stauffer and W. Grimson, *Adaptive background mixture models for real-time tracking*, CVPR99, Fort Collins, CO, 1999.
- [6] A. Prati, I. Mikic, C. Grana, M.M. Trivedi, *Shadow detection algorithms for traffic flow analysis: a comparative study*, in Proceedings of IEEE Intelligent Transportation System Conference (ITSC 2001), Oakland, CA, USA, pp. 340-345, Aug., 2001
- [7] I. Grinias and G. Tziritas, *Motion segmentation and tracking using a seeded region growing method*, EUSIPCO 98, Rhodes, Greece, September 1998.
- [8] D. Izquierdo and Y. Berthoumieu, *Region level segmentation based on a derivative approach for video tracking process*, ICIP 02, TA-P4.5, 2002.
- [9] A. Randriantsoa, Y. Berthoumieu, *Optical Flow Estimation Using Forward-backward Constraint Equation*, ICIP 00, TP03.08, 2000.