

Segmentation Multi - Descripteurs de Scènes Autoroutières

D. Izquierdo¹ J. Becerra² Y. Berthoumieu¹ M. Donias²

¹ UMR 5818 IXL – ENSEIRB, ² UMR 5131 ESI/LAP – ENSEIRB

1, Avenue du docteur Albert Schweitzer, BP 99, 33402 Talence cedex - France

Tél. : +33 (0)5 56 84 65 00 ; fax : +33 (0)5 56 37 20 23

{david.izquierdo, javier.becerra, 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 les objets en mouvement de l'information statique de la scène. A cet effet, une nouvelle méthode de génération d'une image de référence adaptative, insensible aux 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 associées 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.

Mots clés

Segmentation Spatio-temporelle, Référence Adaptative, Expectation-Maximization (EM), Segmentation Couleur, Poursuite.

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 résolu 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 reliée 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 explicitons un processus de suivi qui prend en compte les Attributs d'Objet (AO) et qui est supervisé par l'algorithme EM. La section 5 présente quelques exemples expérimentaux obtenus sur des scènes autoroutières réelles.

2 Segmentation temporelle fondée sur une nouvelle approche IRA

La détection des 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 étape initiale possible consiste à réduire l'observation, ou information utile, à l'aide de l'extraction des masques binaires des régions connexes d'après le résultat de la segmentation temporelle. Dans cette optique la bibliographie en vision par ordinateur offre une grande variété de techniques [5,7]. Une d'entre-elles se dégage tout particulièrement : l'Image de Référence Adaptative (IRA) [4]. Dans la littérature, cette technique se fonde sur la formulation suivante:

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

où B représente l'image de référence, I est l'image courante, α représente la mémoire du système ou la capacité du système à répondre aux changements et k 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 (figures 1 et 2). Le tableau 1 résume les différentes caractéristiques relatives au fond et aux passages d'objets. Ces caractéristiques sont extraites à partir du profil de la courbe de luminosité I_p^t .

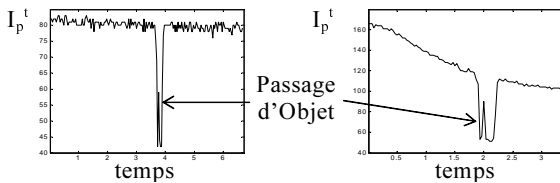


Figure 1 : Avec niveau moyen constant.

Figure 2 : Avec changement de luminosité ambiante.

	Variation Quasi-Linéaire	Luminosité Moyenne
Fond	Lente	Niveau Constant
Objet	Rapide	Forte variabilité

Tableau. 1: Caractéristiques du profil de la courbe de luminosité, I_p^t .

De façon à détecter tous les passages d'objets et à réduire l'influence de 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** qui a en charge d'identifier les passages d'objets. Elle exploite la différence entre les caractéristiques de la courbe de luminosité ambiante et celle des objets en utilisant une approche dérivative, décrite au

paragraphe 2.1. Elle est fondée sur une caractéristique du passages d'un objet: une forte variation du profil de luminosité. L'approche de détection s'appuie sur une approche variationnelle.

- **Deux références auxiliaires** B_1^k et B_2^k (section 2.2) associées à la pente de la courbe de luminosité I_p^t : elles sont introduites pour s'adapter au mieux à une variation 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:

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

$$m_2^{k+1}(p) = m_2^k(p)\beta(p) + (1 - \beta(p))(d^k(p))^2 \quad (4)$$

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

$$\beta(p) = 1 - \alpha D^k(p) \quad (5)$$

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 :

$$B_1^k = B_1^{k-1} \quad (6)$$

$$B_2^k = B_2^{k-1} + L^{k-1}$$

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

$$L^k = \rho L^{k-1} + (1 - \rho)(B_2^k - B_2^{k-1}) \quad (7)$$

La décision finale de la mise à jour de la référence B^k est réalisée par rapport à la minimisation de l'erreur avec la valeur de luminosité actuelle :

$$B^k = \text{Arg min}_{(B_1, B_2)} \left(\left| B_i^k - I^k \right| \right) \text{ pour } i=1,2 \quad (8)$$

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 :

$$B^k = B_1^k = B_2^k = I^k \quad (9)$$

2.3 Processus IRA

Notre implantation s'appuie sur trois images consécutives (du fait du choix de $n=1$ dans l'équation de la variationnelle (éq. 2)). La figure 3 résume le processus IRA au complet. On définit la sortie du processus complet $\mathbf{Ro}^k = \{R_{o_i}^k\}_{i=1, \dots, r}$ 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].

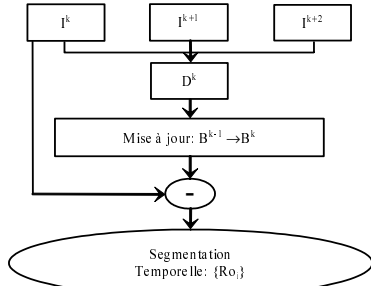


Figure 3 : Diagramme complet de la Segmentation Temporelle.

3 Suppression de l'ombre

Dans le cas d'images en couleur, le passage d'un point éclairé à un point ombré semble se traduire par une chute très nette de la luminosité alors que les valeurs relatives de chrominance rouge et verte sont conservées [6]. Pour classer un point de l'image $I^k(x,y)$ comme étant un point ombré, nous avons implanté un test multicritère fondé sur les deux espaces de couleur, $YCrCb$ et rgb . L'information relative à la luminosité est extraite à partir de la composante Y , et celle relative à la chrominance des composantes Cr , r et g . Un pixel est étiqueté comme ombré dès lors que les quatre équations suivantes sont vérifiées :

$$\begin{aligned} \frac{Y_I(x,y)}{Y_B(x,y)} &< \alpha_Y, \\ |Cr_I(x,y) - Cr_B(x,y)| &< \beta_{Cr}, \\ |r_I(x,y) - r_B(x,y)| &< \beta_r, \\ |g_I(x,y) - g_B(x,y)| &< \beta_g \end{aligned} \quad (10)$$

où α_Y , β_{Cr} , β_r et β_g sont déterminés empiriquement.

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.

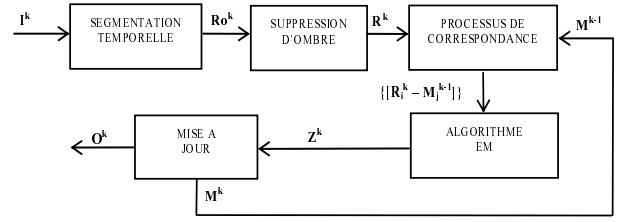


Figure 4 : Diagramme du processus complet de suivi.

4 Processus de suivi

L'objectif du processus de suivi est l'identification de chaque objet en mouvement présent dans la scène. Cette objectif est atteint en plusieurs étapes (figure 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 présent dans la $(k-1)^{\text{ième}}$ image est donné par un ensemble de modèles, notés M_j^{k-1} . Chaque M_j^{k-1} comprend un ensemble d'Attributs d'Objets (AO) permettant l'identification de l'objet dans l'image suivante à partir de l'ensemble des régions R_i^k . Soit $\mathbf{M}^{k-1} = \{M_j^{k-1}\}_{j=1, \dots, q}$ l'ensemble des q modèles associés à la $(k-1)^{\text{ième}}$ image.

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} à travers une compensation de mouvement: $\tilde{M}_j^k = M_j^{k-1}(\theta_j^{k-1})$ où θ_j^{k-1} représente le modèle de mouvement à six paramètres [1,9] correspondant au modèle M_j^{k-1} dans la $(k-1)^{\text{ième}}$ image.

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). Une phase de reconstruction s'appuyant sur une approche EM (section 4.1) se révèle donc nécessaire. Celle-ci conduit à une nouvelle segmentation sous la forme de structures ou « Zones » Z^k recouvrant chacune un objet dans sa totalité. Ce jeu $\{Z^k\}$ a un niveau de signification équivalent à l'objet. Grâce à cet ensemble $\{Z^k\}$, le processus de mise à jour est alors réalisé sur la totalité des Attribut d'Objets de chaque modèle.

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, chaque pixel a une probabilité (à posteriori) d'appartenance à chaque classe dont le modèle statistique de la fonction de densité de probabilité de la valeur actuelle du pixel, I_p^k , conditionnée à un des modèles de mouvement, est caractérisé par une distribution gaussienne conditionnellement au mouvement. La segmentation spatiale est obtenue par association des

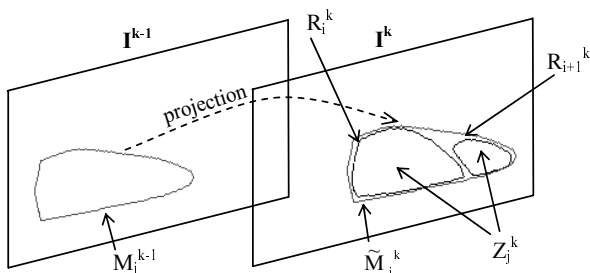


Figure 5 : Modèle : M_j^{k-1} , Modèle projeté : \tilde{M}_j^k , Régions : R_i^k, R_{i+1}^k et la Zone, Z_j^k qui correspond à l'objet réel.

pixels qui ont une probabilité maximum par rapport au modèle de mouvement.

Le résultat extrait de l'algorithme correspond à l'extraction d'une seule zone Z_j^k qui correspond avec le modèle M_j^{k-1} et qui contient une signification complète d'objet (voir figure 5).

Les attributs de chaque modèle sont mis à jour après cet étape du processus EM. Nous utilisons notamment dans notre approche pour estimer les paramètres du modèle de mouvement un estimé de type LMS incrémentale [1,9].

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'occultation 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 d'un modèle colorimétrique.

6 Conclusion

Dans cet article, nous avons proposé une chaîne complète de traitement pour le suivi de véhicules dans des séquences autoroutières. Les atouts de notre schéma sont sa robustesse aux changements de luminosité moyenne de l'image, sa capacité à gérer les occultations et les « sur-segmentations » ou « sous-segmentations » des objets ainsi que la suppression de l'ombre d'un objet. La première partie du schéma présenté est fondé sur une procédure IRA non supervisée qui extrait les régions en mouvement. Un processus de suppression d'ombre a été introduit afin d'améliorer la localisation de l'objet. A partir du résultat binaire de cette première phase de segmentation, une procédure de discrimination est développée 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 une modélisation géométrique et une identification fondée mouvement propre à chaque véhicule. La mise à jour de l'ensemble de leurs descripteurs est réalisée par une procédure optimale s'appuyant sur un algorithme EM.

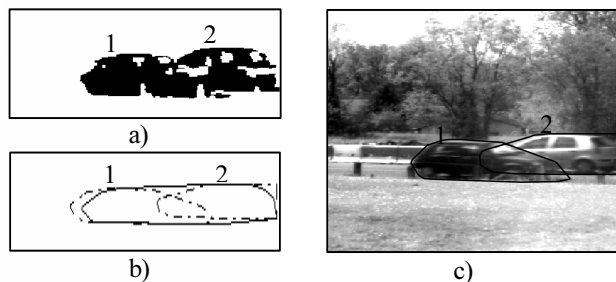


Figure 6 : Image réelle avec traitement d'occultation.

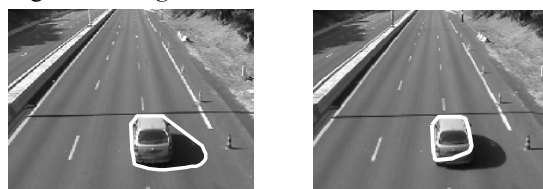


Figure 7 : Suivi avant et après suppression de l'ombre.

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.