

Détection du mouvement fondée sur un modèle d'interaction électrique

Eric Deléchelle

Laboratoire d'Etude et de Recherche en
Instrumentation Signaux et Systèmes
Université Paris 12 Val-de-Marne, Créteil, France
email : delechelle@univ-paris12.fr

Jacques Lemoine

Laboratoire d'Etude et de Recherche en
Instrumentation Signaux et Systèmes
Université Paris 12 Val-de-Marne, Créteil, France
email : lemoine@univ-paris12.fr

Résumé

La méthode présentée a pour but la détection du mouvement dans une image sans estimation du déplacement. La caméra est supposée fixe. La détection est obtenue en utilisant, en chaque point de l'image, une mesure de vraisemblance calculée à partir des valeurs d'une différentiation temporelle suivie d'une intégration spatio-temporelle. L'intégration spatiale est réalisée à l'aide d'un modèle d'interaction de charges électriques distribuées dans le plan de la différence de deux images. La modélisation repose sur l'analogie entre la vraisemblance qu'un point a d'appartenir à une région en mouvement et le niveau de potentiel électrique imposé en ce point par les charges voisines. La prise en compte du contexte temporel est réalisée par extension du modèle 2D à une distribution « volumique » de charges définie sur une tranche de différences consécutives. Les potentiels estimés dans un volume spatio-temporel sont associés à des mesures de vraisemblance qui permettent de définir les régions concernées par un déplacement. Une mesure intermédiaire, appelée Minimum de Vraisemblance et estimée entre deux différences, est définie pour détecter les masques complets des objets mobiles.

Mots clés : Détection du mouvement, potentiel électrique, mesure de vraisemblance, intégration temporelle.

Abstract

Works presented in this paper focuses on the problem of motion detection. The acquisition of images is supposed to have been realised from a static camera. An elaborate algorithm around an electrical interaction analogy allows us to define a measure of likelihood that a pixel to belong to a moving region. A spatio-temporal volume, define by several consecutive image differences, has behaved us naturally to spread the initial 2D model to a volumic distribution of loads. The motion detection is finds some improved so concerning the definition of masks that the elimination of noise.

Key words : Motion detection, electrical potential, likelihood measure, temporal integration.

1. Introduction

Dans un grand nombre de méthodes de détection, la différentiation temporelle permet, dans le meilleur des cas, l'obtention d'une image des contours des objets en mouvement. Les images sont généralement bruitées et il est difficile d'y extraire les vrais contours mobiles par une méthode de seuillage. De plus, les contours détectés ne permettent généralement pas d'obtenir le masque des objets en mouvement sans y introduire des manques. Pour augmenter la robustesse au bruit, ou pallier le manque d'information, il est nécessaire de prendre en compte plus de deux images. Dans ce sens, l'information apportée par les variations temporelles peut faire l'objet d'une intégration temporelle. Nous proposons une méthode simple de détection du mouvement qui repose sur un modèle d'interactions de charges électriques. Cette modélisation permet d'établir, en un point, une mesure de vraisemblance de mouvement en fonction des observations détenues sur un voisinage spatio-temporel donné. La distribution initiale de charges électriques est définie à l'aide d'une probabilité locale de mouvement exprimée à partir d'une différence d'images.

1.1 Travaux antérieurs

De nombreux travaux ont proposé des modèles adaptés au problème de la segmentation d'images en régions statiques et mobiles. Les techniques élaborées dans ce sens sont basées soit sur des heuristiques, soit sur l'utilisation d'une image de référence, ou bien encore sur une procédure d'étiquetage probabiliste. Par exemple, Karmann et Brandt construisent dynamiquement, par filtrage de Kalmann, une image de référence du fond fixe [4], cette procédure permet de tenir compte des variations lentes du fond. Dans les approches probabilistes, telles que celles faisant appel au champs de Markov, l'idée principale est de trouver la configuration la plus probable d'étiquettes à partir des observations. A cet effet, une fonction d'énergie est minimisée, par exemple, à l'aide d'un algorithme de type ICM (Iterated Conditional Modes) [2] [5].

1.2 Probabilité de mouvement

L'idée est d'obtenir une estimation grossière des points jugés très probablement mobiles. De là, une carte binaire des sites mobiles est dressée et permet d'initialiser à l'algorithme présenté. Malgré le danger quelle procure, la différence de deux images consécutives est néanmoins utilisée de manière à exprimer, en chaque point, une probabilité initiale d'appartenance à un contour mobile.

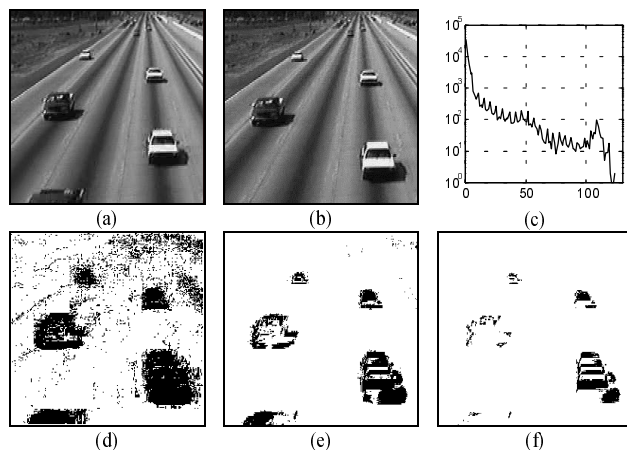


Figure 1 : Seuillage de la différence d'images. (a)-(c) deux images consécutives et histogramme de la différence (en valeur absolue). (d)-(f) résultats de la binarisation pour des valeurs du seuil respectivement égales à 5, 20 et 40.

Lors de l'étude d'une séquence d'images, la discrimination des points mobiles par rapport aux points statiques par seuillage brutal de la différence d'images est un problème difficile. La figure 1 donne le résultat du seuillage de la différence simple de deux images (fig.1.a et b) pour trois valeurs de seuil (fig.1.d à f). L'histogramme de la valeur absolue de la différence est également reporté (fig.1.c). Le choix du seuil est critique dans le sens où une trop faible valeur de ce dernier engendre la détection du bruit (fig.1.d), une valeur élevée permet d'éliminer les fausses détections mais affecte la définition des contours mobiles (fig.1.f). L'étude de l'histogramme permet de fixer manuellement une valeur correcte du seuil de manière à conserver les points des contours mobiles tout en éliminant une grande partie du bruit (fig.1.e). Plutôt que de chercher à trouver la valeur optimale du seuil, il semble intéressant d'estimer, dans un premier temps, la probabilité d'appartenance d'un point à une région en mouvement.

La méthode développée dans ce sens est fondée sur la remarque suivante [8] : les statistiques des dérivées temporelles des images présentent une loi de distribution laplacienne, pour les points des contours mobiles, et gaussienne, dans les régions statiques. On note D_t l'opérateur de différentiation temporelle considéré. Cette opérateur présente une distribution de probabilité laplacienne pour les contours, soit

$$f(D_t(x, y) | c) \approx \frac{1}{2\lambda} \exp\left(-\frac{|D_t(x, y)|}{\lambda}\right)$$

De même, pour les régions statiques, l'opérateur D_t présente une distribution gaussienne

$$f(D_t(x, y) | b) \approx \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(D_t(x, y))^2}{2\sigma^2}\right)$$

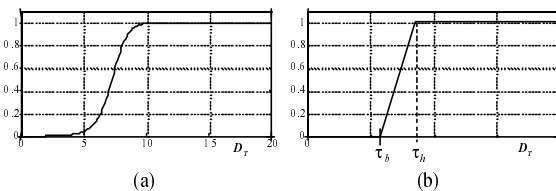


Figure 2 : Probabilité d'appartenance à un contour mobile. (a) courbe obtenue pour des valeurs expérimentales de α , β et γ . (b) approximation linéaire par morceau décrite par deux paramètres τ_b et τ_h .

Nous pouvons exprimer la probabilité conditionnelle $p(\mathbf{p} \in \mathbf{C} | D_t)$ pour qu'un point \mathbf{p} de l'image soit un élément c de l'ensemble \mathbf{C} des contours connaissant la valeur de D_t en ce point. L'expression de $p(c | D_t)$, obtenue à partir du théorème de Bayes, est donnée par

$$p(c | D_t) = \frac{1}{1 + e^{\gamma \frac{(|D_t| - \alpha)^2}{\beta}}}$$

avec $\alpha = \frac{\sigma^2}{\lambda}$, $\beta = \sqrt{2}\sigma$ et $\gamma = \frac{\sigma^2}{2\lambda^2} + \log\left(\frac{p(b)}{p(c)} \frac{2\lambda}{\sqrt{2\pi}\sigma}\right)$,

où $p(b)$ et $p(c)$ sont respectivement les probabilités a priori pour le fond statique et les contours mobiles. La courbe théorique de la fonction de probabilité $p(c | D_t)$ est illustrée sur la figure 2.a pour des valeurs expérimentales des paramètres α , β et γ . Une approximation linéaire par morceau est réalisée (figure 2.b) en définissant des seuils de transition bas et haut, τ_b et τ_h , de la probabilité d'appartenance d'un point à un contour en mouvement. Cette approximation permet de réduire le nombre de paramètres à fixer pour estimer la probabilité conditionnelle. L'expression initiale est alors remplacée par

$$p(c | D_t) \approx \begin{cases} 0, & \text{si } |D_t| \leq \tau_b \\ \frac{|D_t| - \tau_b}{\tau_h - \tau_b}, & \text{si } \tau_b \leq |D_t| \leq \tau_h \\ 1, & \text{si } |D_t| \geq \tau_h \end{cases}$$

En terme de segmentation en régions statiques ou mobiles, la partition de l'image demande l'attribution d'une étiquette à chaque point de l'image en fonction de la probabilité $p(c | D_t)$. De cette manière, et quel que soit le seuil de probabilité retenu, un certain nombre de points seront inévitablement mal classés. Cette remarque vient du fait que la mesure ne dépend que d'un attribut local D_t mesuré en chaque point à partir de deux images. Dans ce sens, la mesure en un point est localisée dans le temps et l'espace et il est nécessaire de procéder à un réajustement des probabilités initiales en fonction du contexte à la fois spatial et temporel.

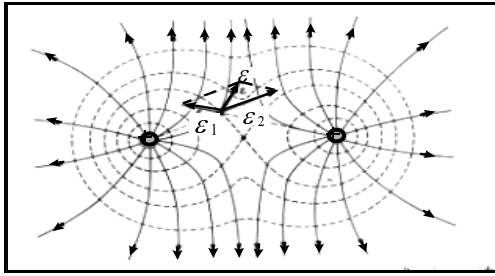


Figure 3 : Champ et interaction électrique. Lignes de champ (trait plein) et courbes équipotentielles (trait pointillé) engendrés par l'interaction de deux charges positives.

2. Le modèle électrostatique

On espère en général améliorer les résultats de la procédure de segmentation en prenant en compte le voisinage de chaque point et les probabilités d'appartenance dans ce voisinage. Cette approche demande d'analyser la compatibilité entre les différentes affectations de points voisins [7]. La probabilité initiale en un point peut être augmentée ou diminuée selon le degré de compatibilité avec la décision prise dans son entourage. Cette approche présente deux difficultés essentielles que sont le choix de la dimension du voisinage et la détermination des coefficients de compatibilité dont le nombre et les valeurs dépendent respectivement du voisinage utilisé et du contenu de l'image. Nous proposons une méthode de réajustement des probabilités initiales qui n'utilise pas de manière explicite de coefficients de compatibilité sur un voisinage donné. La prise en compte du contexte spatio-temporel est réalisée à l'aide d'un modèle d'interactions électriques qui modélisent les influences réciproques entre des points voisins en fonction d'une part de leur distance respectives et d'autre part des probabilités locales $p(c | D_t)$ de mouvement. Une analogie est faite entre la mesure de potentiel V calculée en un point et la vraisemblance qu'a ce point d'appartenir à une région en mouvement.

2.1 Distribution surfacique de charges

La figure 3 illustre le champ électrique \mathbf{E} produit par l'interaction de deux charges de même signe. L'équation de Poisson permet d'exprimer la divergence de \mathbf{E} en fonction de la densité de charges électriques ρ

$$\operatorname{div} \mathbf{E} = \frac{\rho}{\varepsilon_0} = -\nabla^2 V$$

ε_0 est la permittivité du vide et V le potentiel électrique. Dans l'espace libre, où il n'y a pas de charges ($\rho = 0$), l'expression précédente permet d'écrire l'équation de Laplace pour les potentiels

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0$$

Connaissant les points pour lesquels le potentiel est imposé par la présence d'une charge, il est facile de calculer les potentiels sur les sites libres à partir d'une méthode itérative de type Gauss-Seidel [3].

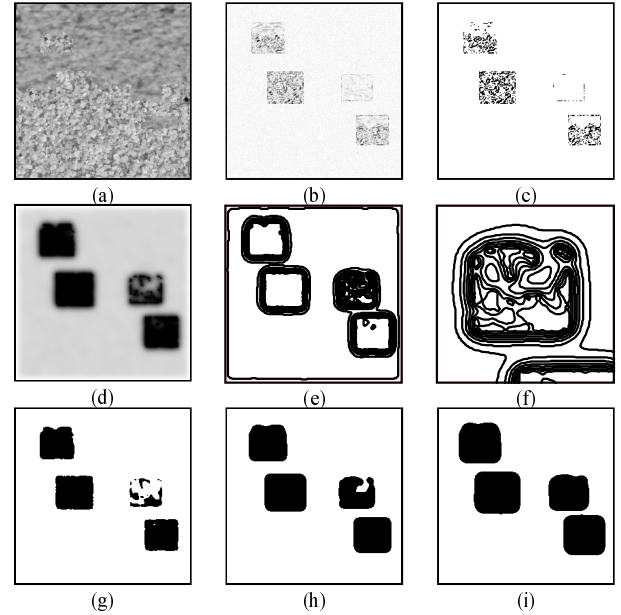


Figure 4 : Potentiels estimés à partir de deux images. (a)-(c) première image, différence, seuillage à $\tau_h = 20$. (d)-(f) potentiels, lignes équipotentielles, détail. (g)-(i) détections pour V supérieur à 0,9, 0,5 et 0,3.

Les sites chargés correspondent aux points pour lesquels la probabilité conditionnelle $p(c | D_t) \equiv V$ est égale à 1, soit pour $|D_t| \geq \tau_h$. Lors de la résolution de l'équation de Laplace,

ces sites ne voient pas leur niveau de potentiel modifié d'une itération à l'autre. Pour les sites libres, le potentiel est initialisé à l'aide de la probabilité locale. Afin de ne pas imposer à certains points le caractère « statique », la valeur du seuil bas τ_h est mise à zéro. Cela favorise le recouvrement de masques complets d'objets mobiles. Un exemple de calcul de potentiels engendrés par une distribution surfacique de charges est donné sur la figure 4. Cet exemple utilise deux images synthétiques sur lesquelles quatre rectangles texturés se déplacent sur un fond qui présente la même texture (fig. 4.a). Les emplacements des charges sont définis par la carte binaire issue du seuillage de la différence des deux images (fig. 4.b). La valeur du seuil, $\tau_h = 20$, est choisie volontairement haute de manière à s'affranchir du bruit.

Ce choix ne permet pas de conserver la totalité de l'information pertinente quant à la détection de certains rectangles mobiles (fig. 4.c). L'estimation des potentiels (fig. 4.d) est obtenue après quelques itérations (< 10). Les lignes équipotentielles sont également données (fig. 4.e et f). La détection des objets mobiles est réalisée en fixant un niveau de potentiel (ou de vraisemblance) minimum pour le seuillage final de l'image des potentiels. Les détections obtenues pour différents niveaux de vraisemblance V sont également illustrés (fig. 4.g à i).

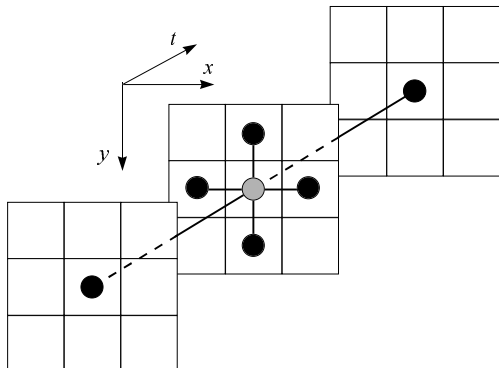


Figure 5 : Voisinage spatio-temporel défini pour la résolution de l'équation de Laplace. Ce voisinage est utilisé dans le cas d'une distribution volumique de charges électriques.

2.2 Modèle volumique ou spatio-temporel

La présence de bruit ne nous permet pas de définir une valeur du paramètre τ_h qui soit adaptée au contenu des images. Dans ce cas et sans a priori sur le niveau de bruit et la nature des objets mobiles, la valeur de ce paramètre doit être fixée arbitrairement haute de manière à éliminer les fausses détections au risque de perdre certaines régions ou parties de régions mobiles. La prise en compte d'un nombre plus important d'images est un atout majeur pour la détection des objets mobiles. Le bruit engendre de fausses détections qui sont généralement localisées dans l'espace et dans le temps. La détection d'une région en mouvement, à un instant donné, peut

être confortée par les observations obtenues aux instants passés et futurs. Nous proposons, dans cette section d'étendre le modèle électrique 2D à un modèle 3D. Cette augmentation de dimension permet d'intégrer l'information temporelle sur un horizon de plus de deux images. L'extension du modèle 2D, à une distribution volumique de charges, est directe. La résolution de l'équation de Laplace est obtenue dans le volume spatio-temporel défini par une tranche de différences d'images consécutives. L'équation initiale est alors remplacée par

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial t^2} = 0$$

où t représente la variable temporelle rapportée à une dimension spatiale. La résolution de cette équation est obtenue en considérant le voisinage spatio-temporel, illustré sur la figure 5, défini par les quatre plus proches voisins spatiaux et les deux plus proches voisins temporels. Comme dans le cas 2D, l'estimation des potentiels est obtenue par une procédure itérative simple dont les différentes étapes sont décrites dans la section suivante.

3. Potentiel et mesure de vraisemblance

Par souci de simplification, les différentes étapes de l'algorithme sont données pour un horizon temporel h défini sur trois images de différences consécutives. La valeur de l'horizon h limite la taille de la fenêtre d'analyse sur l'axe temporel. Pour de longues séquences, cette limite peut être étendue, mais cela augmente la complexité et le temps de réponse. En général, une valeur raisonnable du paramètre h est comprise entre 3 et 5 différences, l'analyse est alors réalisée par fenêtre glissante le long de l'axe temporel.

3.1 Algorithme initial

Nous donnons ici les différentes étapes de l'algorithme 3D de détection. On considère trois différences d'images consécutives calculées aux instants $t-1$, t et $t+1$. Si le triplet $\{t-1, t, t+1\}$ constitue le premier triplet analysé, les opérations données dans les deux premières étapes, à l'instant $t+1$, sont réalisées de la même manière aux instants $t-1$ et t . Dans le cas contraire, les mesures de vraisemblance estimées aux instants $t-1$ et t sur le triplet précédent servent de valeurs initiales pour l'analyse effectuée sur le triplet courant.

- Calcul de $p(c | D_{t+1})$ à l'instant $t+1$.
- Initialisation des mesures $V_0(x, y, t+1)$, on pose $V_0 = 0$ sur les bords de l'image et $V_0 = p(c | D_{t+1})$ pour le reste de l'image. On pose $k = 0$.

- Itération $k+1$, les valeurs V_{k+1} sont modifiées suivant l'équation 1. Les points pour lesquels la mesure de vraisemblance initiale est maximale ($V_0 = 1$) ne sont pas modifiés.
- Si la convergence est atteinte : $\forall \{x, y\}, |V_{k+1} - V_k| < \varepsilon$ alors la procédure est appliquée sur le triplet suivant, sinon on retourne à l'étape 3.

Les mesures notées $V(x, y, t-2)$ et $V(x, y, t+2)$ ne sont pas forcément connues et peuvent être définies soit par prolongation ou fixées a priori. Dans le cas d'une analyse glissante la mesure $V(x, y, t-2)$ correspond à l'estimation obtenue sur le triplet précédent, sa valeur n'étant plus modifiée.

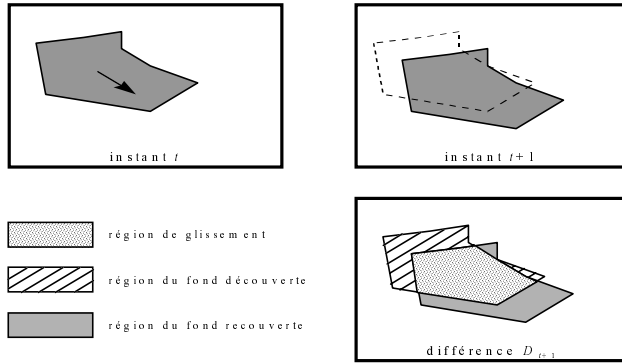


Figure 6 : Régions définies par la différence de deux images.

3.2 Minimum de vraisemblance

La transposition directe du modèle surfacique de charges électriques à une distribution volumique, nous donne le moyen d'intégrer l'information disponible sur un laps de temps plus important. De cette manière, l'intégration temporelle doit permettre d'éliminer un grand nombre de fausses détections dues au bruit. Mais il n'en reste pas moins que les mesures de vraisemblance obtenues ne permettent pas, à proprement parlé, d'extraire les masques des objets mobiles. En effet, ces mesures traduisent, dans le meilleur des cas, la présence d'un changement dû à un objet en mouvement. La région ainsi définie, par seuillage de l'image des vraisemblances, est non seulement constituée du support de l'objet mobile mais également de la zone du fond découverte par cet objet d'un

instant à l'autre. La figure 6 illustre les différentes régions ainsi définies par la différence entre deux images consécutives. Cette différence fait apparaître trois régions. La première correspond à la zone de glissement de l'objet sur lui-même. Cette zone peut être absente si l'amplitude du déplacement est trop grande. Les deux autres régions sont définies par les parties du fond qui sont respectivement découvertes et recouvertes par l'objet mobile.

Ainsi, l'utilisation d'une différence simple ne permet pas la discrimination entre ces régions. Dans un grand nombre de cas, cette discrimination peut être faite à partir de trois images. Par exemple, Orkisz propose de détecter les contours en mouvement par une combinaison des gradients spatiaux de trois images consécutives [6]. Nous proposons une modification de l'algorithme 3D initial de manière à introduire une mesure qui porte uniquement sur la détection du masque des objets.

3.3 Algorithme modifié

La modification apportée permet d'éliminer la région du fond découverte par l'objet mobile. Pour cela, on définit un plan intermédiaire P entre deux différences. Ce sur-échantillonnage du volume spatio-temporel permet de définir une mesure de vraisemblance VM intermédiaire (ou moyenne). Cette mesure, obtenue par moyennage spatio-temporel, doit présenter une valeur maximale à l'emplacement de l'objet. Cette remarque est illustrée sur la figure 7.

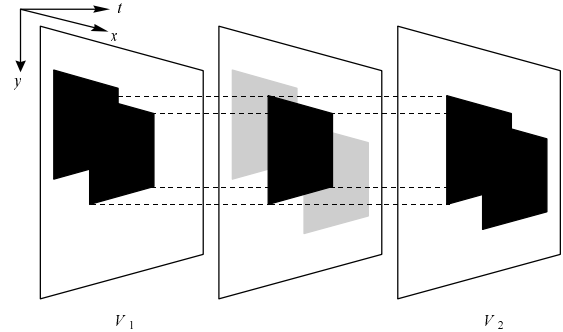


Figure 7 : Interprétation du plan intermédiaire. Les mesures minimales et maximales apparaissent respectivement en blanc et noir. Les régions grisées sont celles du fond recouvert et découvert par l'objet.

(équation 1)

$$V_{k+1}(x, y, t-1) = \frac{1}{6} [V(x, y, t-2) + V_{k+1}(x-1, y, t-1) + V_{k+1}(x, y-1, t-1) + V_k(x+1, y, t-1) + V_k(x, y+1, t-1) + V_k(x, y, t)]$$

$$V_{k+1}(x, y, t) = \frac{1}{6} [V_{k+1}(x, y, t-1) + V_{k+1}(x-1, y, t) + V_{k+1}(x, y-1, t) + V_k(x+1, y, t) + V_k(x, y+1, t) + V_k(x, y, t+1)]$$

$$V_{k+1}(x, y, t+1) = \frac{1}{6} [V_{k+1}(x, y, t) + V_{k+1}(x-1, y, t+1) + V_{k+1}(x, y-1, t+1) + V_k(x+1, y, t+1) + V_k(x, y+1, t+1) + V(x, y, t+2)]$$

Expérimentalement, le moyennage spatio-temporel a pour conséquence un lissage trop important des mesures VM , et leur faible variation rend la détection très sensible à la valeur du seuil. Pour cela, la mesure VM est remplacée par le Minimum des mesures de Vraisemblance, noté MV , qui est simplement défini comme étant le minimum entre deux mesures de vraisemblance consécutives. La comparaison entre les résultats obtenus par les deux mesures est donnée sur la figure 8.

L'algorithme initial est modifié de la manière suivante. On se donne trois images consécutives qui permettent d'exprimer deux différences d'images, D_{t-1} et D_t . Un plan intermédiaire P_t , à valeurs MV initialement nulles, est défini entre les deux différences temporelles. Le triplet défini par les images $\{D_{t-1}, P_t, D_t\}$ est analysé de la manière suivante :

- Calcul de $p(c|D_t)$ à l'instant t .
- Initialisation des mesures $V_0(x, y, t)$, on pose $V_0 = 0$ sur les bords de l'image et $V_0 = p(c|D_t)$ pour le reste de l'image. La mesure $V_0(x, y, t-1)$ est initialisée à l'aide de la vraisemblance estimée sur le triplet précédent. On pose $k = 0$.
- Itération $k + 1$, les valeurs V_{k+1} sont modifiées suivant l'équation 2. Les points pour lesquels la mesure de vraisemblance initiale est maximale ($V_0 = 1$) ne sont pas modifiés, les MV sont toujours modifiés.
- Si la convergence est atteinte: $\forall \{x, y\}, |MV_{k+1} - MV_k| < \varepsilon$ alors la procédure est appliquée sur le triplet suivant, sinon on retourne à l'étape 3.

L'algorithme est facilement adapté pour un horizon temporel plus important. En l'occurrence, dans le cas de quatre images consécutives, trois différences d'images deux plans intermédiaires sont utilisés. Les équations de la troisième étape sont complétées de manière à tenir compte de l'augmentation de l'horizon temporel. La figure 9 donne l'implantation de l'algorithme pour un horizon temporel de quatre images.

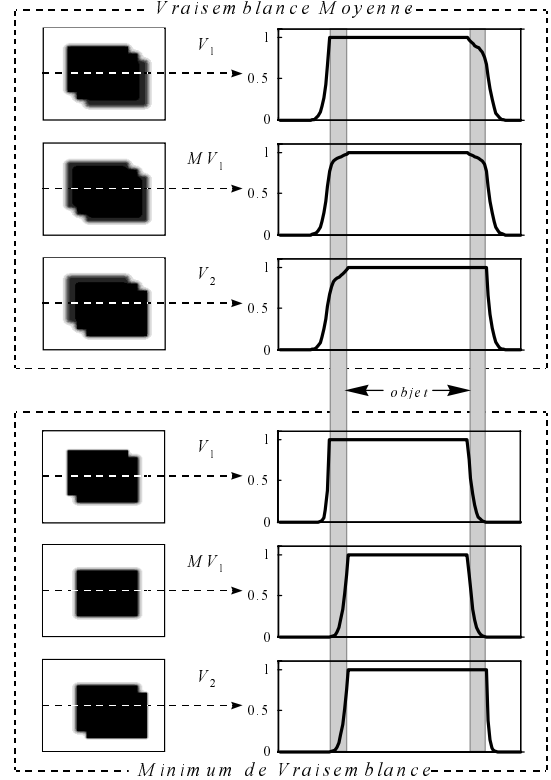


Figure 8: Vraisemblance Moyenne et Minimum de Vraisemblance. La comparaison est réalisée pour la configuration illustrée sur la figure 7.

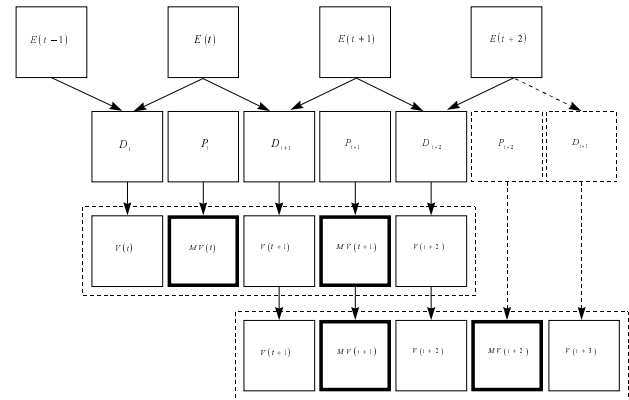


Figure 9: Implantation de l'algorithme 3D pour un horizon temporel de quatre images. Trois différences et deux plans intermédiaires sont analysés simultanément. La première analyse permet d'initialiser en partie les mesures définies lors de la seconde analyse.

(équation 2)

$$\begin{aligned}
 V_{k+1}(x, y, t-1) &= \frac{1}{6} [V(x, y, t-2) + V_{k+1}(x-1, y, t-1) + V_{k+1}(x, y-1, t-1) + V_k(x+1, y, t-1) + V_k(x, y+1, t-1) + MV_k(x, y, t)] \\
 MV_{k+1}(x, y, t) &= \min\{V_{k+1}(x, y, t-1), V_k(x, y, t)\} \\
 V_{k+1}(x, y, t+1) &= \frac{1}{6} [MV_{k+1}(x, y, t) + V_{k+1}(x-1, y, t+1) + V_{k+1}(x, y-1, t+1) + V_k(x+1, y, t+1) + V_k(x, y+1, t+1) + V(x, y, t+2)]
 \end{aligned}$$

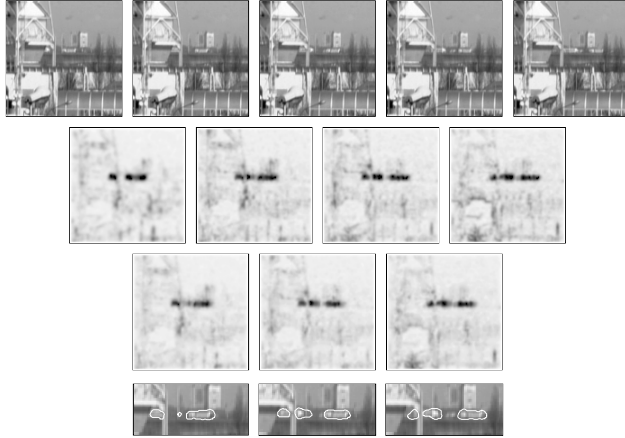


Figure 10 : Exemple de détection. Première ligne : cinq premières images de la séquence. Seconde ligne : mesures de vraisemblance ($t = 2, \dots, 5$). Troisième ligne : MV associés ($t = 2, 3, 4$). Quatrième ligne : contours détectés.

4. Résultats et discussion

Le premier exemple porte sur une séquence sur laquelle on distingue deux véhicules en mouvement. La caméra, de type CCD, est tenue à la main ce qui engendre certaines oscillations. L'analyse porte sur une tranche de cinq images, l'horizon temporel est défini sur trois images consécutives. Les difficultés principales rencontrées dans cet exemple sont, un niveau de bruit important, la taille des objets est de quelques points, les déplacements des véhicules sont de l'ordre de la moitié de leur longueur. Le résultat de la détection est donné sur la figure 10. La valeur du paramètre τ_h est prise égale à 5. Les mesures de vraisemblance V sont données sur la seconde ligne de la figure. Les détections aux instants $t = 2$ à 4 sont illustrées sur les images en niveaux de gris (dernière ligne) pour une valeur des MV supérieures à 0,5. Pour illustrer l'apport de l'intégration temporelle, nous avons appliqué l'algorithme 2D (section 2.1) entre les instants 4 et 5 de cette séquence. Le résultat de la détection est illustré sur la figure 11. La valeur de τ_h est la même que précédemment.

Le second exemple porte sur une séquence d'intérieur dans laquelle un individu se déplace dans une pièce, figure 12. Le recouvrement du personnage sur lui même est très faible, voire inexistant d'une image à l'autre. De plus, sa taille apparente est relativement importante par rapport à celle de l'image. De nombreuses régions qui le composent sont peu contrastées par rapport au fond des images. La détection est illustrée sur la figure 12 pour deux valeurs du paramètre τ_h . Dans le premier cas, $\tau_h = 5$, le seuillage des différences d'images fait apparaître de nombreuses régions de bruit. Malgré cela, l'analyse par quadruplet d'images ($h = 4$) permet d'éliminer une grande partie de ce bruit en particulier en dehors des régions de recouvrement. Dans le second cas, l'élimination

complète du bruit est obtenue en augmentant la valeur du seuil initial à $\tau_h = 30$. Seulement, certaines régions peu texturées et faiblement contrastées du personnage ne sont pas détectées. Le résultat de la détection est donné à la figure 13.

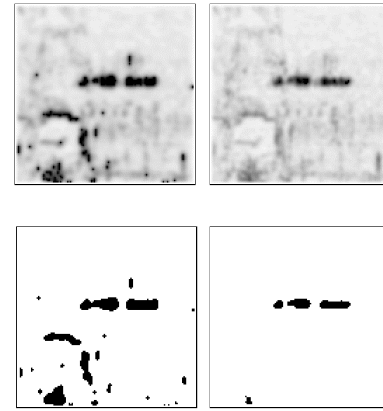


Figure 11 : Détections avec et sans intégration temporelle.

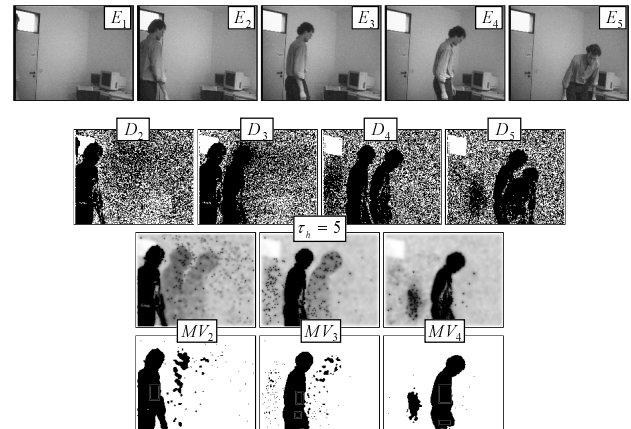


Figure 12 : Détections sur la séquence "Bureau". L'analyse est réalisée par quadruplet d'images. La valeur du seuil est égale à $\tau_h = 5$.

5. Conclusion

Nous avons présenté une méthode originale de détection d'objets mobiles sur une séquence d'images. Un modèle d'interactions de charges électriques, définies dans le volume spatio-temporel représenté par plusieurs différences d'images consécutives, nous a permis d'élaborer une procédure de réajustement de probabilités de mouvement initiales calculées à partir d'une simple différence d'images. Le caractère spatio-temporel de l'analyse, naturellement détenu par le modèle, est un avantage certain quant à l'analyse de phénomènes temporels tels que le mouvement. De plus, la complexité de calculs engendrée par cette approche itérative est très réduite, et la convergence obtenue en quelques itérations (< 10).

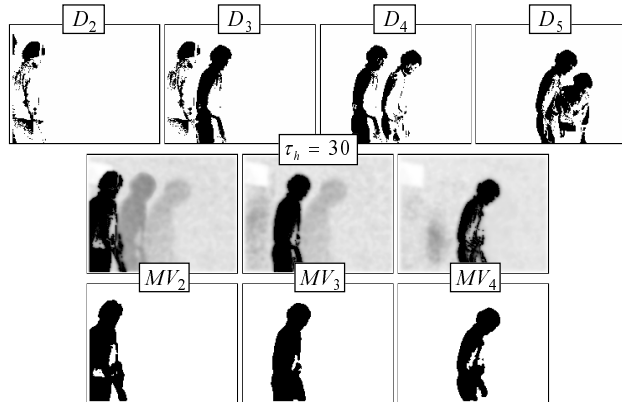


Figure 13 : Detections sur la séquence "Bureau" ($\tau_h = 30$). Les résultats peuvent être comparés à ceux obtenus pour $\tau_h = 5$, illustrés sur la figure 12.

On note également que cette approche ne demande la détermination que d'un seul paramètre, τ_h .

Le résultat de la figure 13 montre la limite de l'algorithme proposé dans son aptitude à recouvrir les masques complets d'objets mobiles de grandes tailles. Ce problème général à toute procédure de détection est généralement surmonté en intégrant l'algorithme développé dans un contexte multirésolution. En particulier, la multirésolution voire spatio-temporelle permet, de s'affranchir non seulement du bruit (filtrage passe-bas des images) mais également de tenir compte (par une analyse progressive) des variations pertinentes dans les régions de mouvement. La difficulté essentielle est de trouver une implantation directe de l'algorithme monorésolution dans une représentation multirésolution. Par exemple, Caplier résout le problème de la détection de grands objets en établissant une version multirésolution spatio-temporelle de l'approche markovienne de détection du mouvement [2].

Nous nous sommes également intéressés à établir une telle implantation dans le cadre du modèle présenté ici. L'idée est de générer une représentation pyramidale laplacienne des images [1]. Cette représentation permet de décomposer l'information sur plusieurs niveaux de résolution, puis de reconstruire partiellement cette information, dans les régions pertinentes, en utilisant les mesures de vraisemblance estimées sur chaque niveau de la pyramide. Les résultats ainsi obtenus [3] sont tout à fait satisfaisants et feront l'objet d'une publication prochaine.

Références

- [1] P.J. Burt, "Fast filter transforms for image processing", *CGIP*, vol. 16, pp. 20-51, 1981.
- [2] A. Caplier, « Modèles markoviens de détection de mouvement dans les séquences d'images : approche spatio-temporelle et mise en œuvre temps réel », *Thèse, TIRF, INPG*, 1995.
- [3] E. Deléchelle, « Analyse de scènes dynamiques : détection du mouvement fondée sur un modèle d'interaction électrique ». *Thèse, LERISS, Université Paris XII*, 1997.
- [4] K.P. Karmann and A. von Brandt, "Detection and tracking of moving objects by adaptive background extraction", *IAPR, 6th Conf. on Image Analysis*, vol. 2, pp. 1051-1058, 1989.
- [5] F. Luthon and A. Caplier, "Motion detection and segmentation in image sequences using Markov Random Field modelling", *4th Eurographics Animation and Simulation Workshop*, Barcelona, pp. 265-275, 1993.
- [6] M. Orkisz, « Localisation d'objets mobiles dans les scènes naturelles filmées par une caméra fixe ». *Traitement du signal*, vol. 9, n° 4, pages 325-345, 1992.
- [7] J.G. Postaire, « De l'image à la décision, analyse des images numériques et théorie de la décision », *Dunod Info.*, 1987.
- [8] K.P. Pratt, "Digital Image processing", *Wiley-Interscience Publication*, 1991.