

# Analyse globale du flux optique pour la détection d'évènements dans une scène de foule

Yassine Benabbas\*, Nacim Ihaddadene\*, Thierry Urruty\*, Chabane Djeraba\*

\* Laboratoire d'Informatique Fondamentale de Lille (LIFL), UMR 8022 USTL/CNRS  
Université des Sciences et Technologies de Lille 1  
59650 Villeneuve d'Ascq, France  
{yassine.benabbas,nacim.ihaddadene,thierry.urruty,chabane.djeraba}@lifl.fr

**Résumé.** Les systèmes de vidéo-surveillance sont de plus en plus autonomes dans la détection des événements anormaux. Cet article présente une méthode de détection des flux majeurs et des événements qui surviennent dans une scène de foule. Ces détections sont effectuées en utilisant un modèle directionnel construit à partir d'un mélange de lois de von Mises appliqué à l'orientation des vecteurs de mouvement. Les flux majeurs sont alors calculés en récupérant les orientations les plus importantes des mélanges. Divers événements se produisant dans une foule sont aussi détectés en utilisant en plus du modèle d'orientation, un modèle probabiliste de magnitude des vecteurs de mouvement. Les résultats de l'expérimentation sur un échantillon de vidéos d'événements sont présentés.

## 1 Introduction

La sécurité des personnes et des biens est un des problèmes majeurs dans les zones publiques telles que les aéroports, les stations de métro, les centres commerciaux ou les places publiques. Le traitement automatique des vidéos provenant de caméras de surveillance est de plus en plus utilisé pour présenter une information pertinente aux opérateurs qui doivent agir dans les situations critiques, dangereuses ou inhabituelles. Ces dernières années ont vu l'intégration dans les systèmes de vidéo-surveillance d'algorithmes de détection de mouvements, d'événements, de bagages abandonnés ou de suivi de personnes. Cependant, vu la complexité du problème, peu de systèmes se sont penchés sur les situations impliquant des foules de personnes.

L'analyse des flux de personnes consiste à détecter les tendances de mouvement dans les zones surveillées. Elle devient nécessaire lorsque le suivi d'objets individuels échoue ; ce qui est souvent le cas dans une scène de foule. Cette analyse est effectuée en traitant l'information de mouvement à travers des images successives. La détection d'événements est définie comme étant la détection des situations qui attirent l'attention d'une personne (Shyu et al., 2008). C'est un domaine vaste du à l'immense quantité d'événements possibles. En plus, la définition d'un événement change d'une personne à une autre et dépend fortement du contexte. Beaucoup d'efforts ont été fournis pour la détection d'événements atomiques qui représentent un élément de base pour la détection de situations plus complexes.

Dans cet article, nous introduisons un modèle probabiliste pour représenter une scène. Ce modèle gère efficacement la complexité des scénarios et l'imprévisibilité des comportements. Notre approche a été appliquée sur une sélection d'évènements et expérimentée sur l'échantillon de vidéos du workshop *PETS'2009*<sup>1</sup>.

Cet article est organisé comme suit. Tout d'abord, La section 2 présente quelques travaux antérieurs traitant du problème de l'analyse de foule. La section 3 détaille la modélisation de la scène. La section 4 décrit la méthodologie adoptée pour l'extraction des flux d'une scène et la détection d'évènements dans une scène de foule. Les résultats de l'expérimentation sur l'échantillon de vidéos *PETS'2009* sont présentés dans la section 5. La section 6 conclue cet article et décrit les travaux futurs potentiels.

## 2 Travaux antérieurs

Les approches traditionnelles pour l'analyse des événements dans une séquence vidéo sont composées des étapes suivantes (Hu et al., 2008) : la détection de tous les sujets mobiles présents dans la scène à un instant  $t$  ; le suivi des sujets détectés dans les instants suivants ; et l'analyse des allures et des trajectoires des sujets pour la détection d'évènements ou d'activités. Cependant, cette approche a beaucoup de faiblesses quand elle est appliquée à des scènes ou scénarios complexes impliquant une large foule. Parallèlement, Les approches globales traitent l'intégralité de la scène sans se focaliser sur les individus séparément. Elles sont classées en deux catégories : la première consiste à estimer la densité de la foule et la seconde à extraire des motifs de mouvement ou détecter des événements dans une scène de foule.

Dans la première catégorie, on distingue les méthodes basées sur l'extraction des textures et sur l'analyse des proportions des surfaces en mouvement (Marana et al., 1997; Rahmalan et al., 2006; Lin et al., 2001; Ma et al., 2004). Ces méthodes fournissent une analyse statique intéressante pour la surveillance des foules, mais ne détectent pas les événements anormaux. Il existe aussi quelques techniques basées sur le flux optique (Boghossian et Velastin, 1999; Davies et al., 1995) qui détectent des foules stationnaires, ou suivent des individus en utilisant plusieurs caméras (Cupillard et al., 2004).

Dans la deuxième catégorie, le but est de détecter les événements anormaux dans une foule en se basant sur les motifs de mouvement. Le principe d'extraction des motifs de mouvement est de modéliser les comportements les plus fréquents et de considérer les événements anormaux comme des cas aberrants. La déviation des comportements typiques est ainsi utilisée pour caractériser l'anormalité.

Plusieurs techniques ont été proposées pour cette catégorie. (Andrade et al., 2006a,b) combinent les modèles de Markov cachés avec l'analyse en composantes principales des vecteurs du flux optique pour détecter des scénarios d'urgences. Cependant, les expérimentations ont été portées sur des données simulées. (Ali et Shah, 2007) utilisent la dynamique des particules Lagrangiennes pour détecter les instabilités du flux. Cette méthode est efficace pour la segmentation des grandes densités de foules (marathons, événements politiques et religieux, etc.). (Ihaddadene et Djeraba, 2008) détectent les situations d'écroulement en se basant sur une mesure qui décrit le degré d'organisation ou de désordre des vecteurs de flux optique. Cette approche fonctionne sur des zones unidirectionnelles (e.g. escalators). (Mehran et al., 2009)

---

1. <http://www.cvg.rdg.ac.uk/PETS2009/>

utilisent le flux optique pure pour détecter les comportements anormaux dans la foule en utilisant un modèle de force sociale. (Wright et Pless, 2005) déterminent les motifs de mouvement persistants avec une distribution jointe globale des distributions indépendantes des gradients de luminosité locaux. Cette variable aléatoire est modélisée avec un mélange Gaussien. Cette approche assume que tous les mouvements sur une image sont cohérents (e.g. voitures), cette hypothèse est violée quand les piétons se déplacent indépendamment.

L'approche proposée contribue à la détection des orientations majeures dans une scène complexe, dû au fait qu'elle construit en-ligne un modèle probabiliste sur l'orientation du mouvement sur la scène qui opère en temps réel. Elle contribue aussi dans la détection d'évènements dans la foule en suivant des groupes de personne au lieu de suivre chaque personne individuellement, ce qui facilite la détection d'évènements se produisant sur des foules.

### 3 Détection et suivi de groupes

L'approche proposée est composée de plusieurs étapes (Figure 1). La première consiste à extraire un ensemble de points d'intérêt dans l'image courante. Ces points sont par la suite traqués dans l'image suivante en utilisant des techniques de calcul du flux optique. Les points statiques sont supprimés pour se focaliser sur les sujets en mouvement. La scène est divisée en blocs, et chaque vecteur de mouvement est attaché au bloc correspondant.

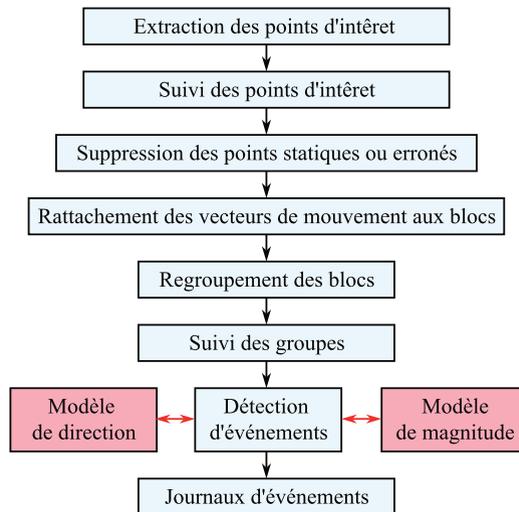


FIG. 1 – Étapes de l'algorithme

L'étape de classification des blocs regroupe les blocs voisins ayant une orientation et une vitesse similaires. Enfin, les évènements sont détectés en exploitant les informations du suivi de groupes, le modèle de magnitudes et le modèle directionnel.

### 3.1 Détection et suivi de points d'intérêt

Ces étapes permettent d'extraire un ensemble de points d'intérêt à partir d'une image de la séquence vidéo en utilisant la méthode décrite dans (Harris et Stephens, 1988). Nous considérons que dans les scènes de vidéo surveillance, les positions des caméras et les conditions de luminosité permettent d'extraire un nombre conséquent de points d'intérêt. Tous les points d'intérêt de l'image sont traqués dans l'image suivante en utilisant l'algorithme de calcul de flux optique (Lucas et Kanade, 1981; Shi et Tomasi, 1994). Le résultat de cette opération à l'instant  $t$  est un ensemble  $V_t$  de vecteurs à 4 dimensions :

$$V_t = \{V_1 \dots V_N | V_i = (X_i, Y_i, A_i, M_i)\}$$

où  $X_i$  et  $Y_i$  sont les coordonnées du point  $i$  par rapport au repère de l'image  $I_t$ ;  $A_i$  est l'angle (ou l'orientation) de mouvement du point  $i$ ;  $M_i$  est la magnitude de mouvement du point  $i$ , elle correspond à la distance entre la position du point  $i$  dans l'image  $I_t$  et sa nouvelle position dans l'image  $I_{t+1}$ . Cette étape permet d'enlever les points d'intérêts statiques qui ont une magnitude de mouvement inférieure à une magnitude minimale.

Finalement, l'image est divisée en blocs de  $B_x$  lignes et  $B_y$  colonnes. Chaque vecteur de mouvement est affecté au bloc adéquat selon son origine. Nous considérons des blocs de dimension  $16 \times 16$  pixels afin d'obtenir un bon équilibre entre le temps de traitement et la qualité des résultats.

### 3.2 Le modèle de magnitude et le modèle de direction

Dans cette étape, on calcule la direction du flux optique dans chaque bloc. Ces directions sont les entrées de la loi de probabilité du bloc. La loi locale pour un bloc  $B_{x,y}$  est construite en utilisant un mélange de lois de von Mises (aussi appelée loi normale circulaire). La probabilité d'un variable angulaire  $\theta$  est donnée par :

$$P_{x,y}(\theta) = \sum_{i=1}^K \omega_{i,x,y} \cdot V(\theta; \mu_{i,x,y}, m_{i,x,y}) \quad (1)$$

où  $K$  est le nombre de distributions. Il représente le nombre maximum d'orientations principales considéré.  $\omega_{i,x,y}$ ,  $\mu_{i,x,y}$  et  $m_{i,x,y}$  sont respectivement le poids, la direction modale et le paramètre de concentration de la  $i^{me}$  loi du bloc  $B_{x,y}$ .  $V(\theta; \mu, m)$  est la loi de von Mises de direction modale  $\mu$  et paramètre de concentration  $m$ . Elle se caractérise par sa densité sur  $[0, 2\pi[$  :

$$V(\theta; \mu, m) = \frac{1}{2\pi I_0(m)} \exp[m \cos(\theta - \mu)] \quad (2)$$

où  $I_0(m)$  est la fonction de Bessel modifiée de première espèce et d'ordre 0 définie par :

$$I_0(m) = \sum_{r=0}^{\infty} \left(\frac{1}{r!}\right)^2 \left(\frac{1}{2}m\right)^{2r} \quad (3)$$

La figure 2 montre une représentation de la fonction densité d'un bloc, on y distingue les deux tendances d'orientation : gauche et droite.

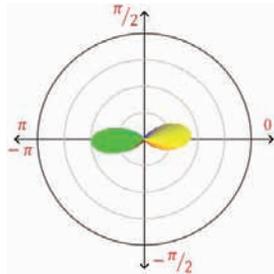


FIG. 2 – Représentation d'un mélange de 2 lois de von Mises

Dans les espaces publics, les personnes se dirigent généralement dans des directions différentes. Ainsi, certaines zones ont des motifs de mouvement qui contiennent 2 orientations principales ou plus. Par exemple, il y a 4 orientations principales dans un passage piéton, 2 orientations opposées pour les piétons et deux autres pour les voitures. Il serait intéressant de savoir qui des piétons ou des voitures emprunte ce passage. Ceci peut être fait en utilisant le modèle directionnel. En effet, le modèle de mélange de chaque bloc peut contenir jusqu'à 4 orientations principales. Le poids nous informe de l'importance d'une orientation.

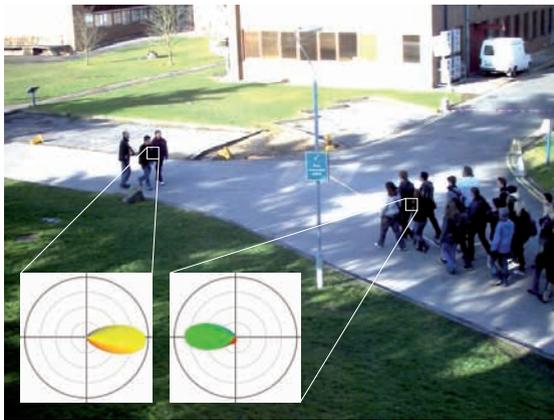


FIG. 3 – Représentation des directions dominantes par bloc

Pour chaque image, les paramètres du mélange sont mis à jour en utilisant un algorithme en-ligne d'approximation E-M (Kaewtrakulpong et Bowden, 2001) qui est à l'origine utilisé pour estimer un mélange Gaussien. L'algorithme a été adapté pour traiter des observations circulaires au lieu d'observations linéaires. L'inverse de la variance  $\sigma$  de la loi normale est considéré comme étant le paramètre de concentration ;  $m = 1/\sigma^2$ . La figure 3 montre la représentation la loi de probabilité pour 2 blocs. Les tendances de direction peuvent être perçues.

Dans ce papier, on choisit  $K = 4$  afin de représenter les 4 points cardinaux.

Le modèle de magnitude est un mélange Gaussien unidimensionnel sur la magnitude moyenne du mouvement entre deux images. Elle est définie comme suit :

$$P(x) = \sum_{k=1}^4 \omega_k \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left(-\frac{(x - \mu_k)^2}{2\sigma_k^2}\right) \quad (4)$$

où  $\omega_k$ ,  $\mu_k$  et  $\sigma_k$  sont respectivement le poids, la moyenne et la variance de la  $k^{me}$  Gaussienne. Les paramètres du mélange sont appris avec un algorithme d'apprentissage en-ligne sur les données vidéo 'S0.normal-flow' appartenant à la base de vidéos de PETS'2009. Ces séquences contiennent des personnes qui marchent.



(a) Image initiale



(b) Tendances de direction

FIG. 4 – Représentation des flux multiples d'une image

### 3.3 Regroupement des blocs

L'objectif de cette opération est de rassembler les blocs d'image entre eux pour obtenir des groupes représentant des entités ayant des directions similaires. Ces groupes peuvent représenter une ou plusieurs personnes. Les critères d'affectation d'un bloc à un groupe sont la direction et la vitesse. Ainsi, les blocs voisins ayant une magnitude moyenne et une orientation principale similaires seront rattachés au même groupe.

Chaque bloc  $B_{x,y}$  est défini par sa position  $P_{x,y} = (x,y); x = \overline{1..Bx}, y = \overline{1..By}$ , et orientation  $\Omega_{x,y} = \mu_{0,x,y}$  (voir section 3.2).

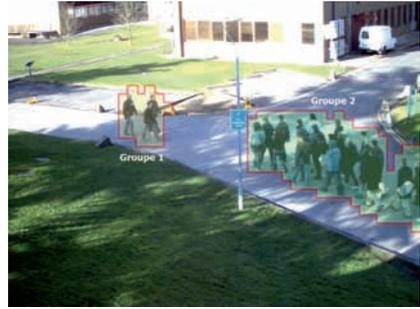
La figure 5 illustre le résultat de cette étape.

### 3.4 Suivi de groupes

Le suivi permet de retracer la trajectoire de la personne ou du groupe de personnes depuis son apparition dans le champ de la caméra jusqu'à sa disparition. Cette opération est effectuée en mettant en correspondance les barycentres des groupes dans l'image  $f$  avec les barycentres de l'image  $f + 1$ . Chaque image  $f$  est définie par ses groupes  $\{C_{1,f}, C_{2,f}, \dots, C_{n_f,f}\}$  où  $n_f$  est le nombre de groupes détectés dans l'image  $f$ . Chaque groupe  $C_{i,f}$  est décrit par son barycentre



(a) Segmentation des blocs



(b) Image d'origine avec les groupes résultants

FIG. 5 – Regroupement des blocs

$O_{i,f}$  et son orientation moyenne  $X_{i,f}$ . Le groupe  $C_{m,f+1}$  correspond au groupe  $C_{i,f}$  si son barycentre est le plus proche de  $C_{i,f}$  sans dépasser une distance minimale. En d'autres termes, elle doit satisfaire ces deux conditions :

$$\left\{ \begin{array}{l} m = \underset{j}{\operatorname{argmin}}(D(O_{i,f}, O_{j,f+1})) \\ \text{et} \\ D(O_{i,f}, O_{m,f+1}) < \tau \end{array} \right. \quad (5)$$

où  $\tau$  est la distance minimale entre deux groupes (on a choisi  $\tau = 5$ ). Si aucune correspondance n'a été trouvée (aucun groupe  $C_{m,f+1}$  ne satisfait les deux conditions), alors le groupe  $C_{i,f}$  a disparu et n'est plus suivi dans les prochaines images.

## 4 Détection d'évènements dans une scène de foule

Dans cette section nous décrivons les traitements nécessaires pour détecter certains événements liés aux foules de personnes. Les scénarios retenus font partie des événements décrits dans le workshop PETS'2009 :

- Évènements liés à la vitesse de mouvement : il s'agit de détecter si les personnes qui forment la foule marchent ou courent. Ces événements peuvent être détectés en exploitant la magnitude des vecteurs de flux optique à travers les images.
- Le rassemblement : cet événement se produit quand deux ou plusieurs groupes se rejoignent pour former un seul groupe.
- La division : cet événement se produit lorsque les personnes qui forment un groupe se séparent pour des raisons telles que la dispersion ou l'évacuation.

Les événements de la première catégorie sont détectés en comparant la magnitude moyenne du flux optique de chaque image au modèle de magnitude de la scène obtenu par apprentissage. Les événements de la deuxième et troisième catégorie sont détectés en analysant la position et la vitesse des groupes. Une explication plus détaillée sera donnée dans ce qui suit.

#### 4.1 Évènements relatifs à l'allure

Comme décrite précédemment, l'idée principale est de comparer la magnitude moyenne des vecteurs de mouvement de chaque image au modèle de magnitude de la scène pour obtenir les probabilités  $P_{march}$ ,  $P_{cour}$  des évènements *marcher* et *courir* respectivement. Dans cet article, puisque les vecteurs de mouvement sont considérés pour la détection bas niveau, et  $P_{cour} = 1 - P_{march}$ . Le calcul de  $P_{cour}$  permettra donc de déduire l'occurrence de l'évènement *courir* et *marcher*. Sous l'hypothèse qu'il y a plus de probabilité pour un groupe de rester dans son état marchant que de passer subitement vers un état différent, alors la probabilité finale des évènements *courir* et *marcher* est une somme pondérée des probabilités précédentes et de la probabilité courante. Formellement, une image  $f$  avec une magnitude moyenne de flux optique  $m_f$  contient un évènement *courir* si :

$$\sum_{l=f-h}^f w_{f-l} \cdot P_{cour}(m_l) > \vartheta_{cour} \quad (6)$$

où  $\vartheta_{cour}$  est le seuil de probabilité l'évènement *courir*.  $h$  est le nombre d'images précédentes à considérer. Chaque état précédent a un poids  $w_l$  (on a défini  $h = 1$ ,  $w_0 = 0.8$  et  $w_1 = 0.2$ ).  $P_{cour}(m_l)$  est la probabilité d'observer  $m_l$ . Elle est obtenue en comparant  $m_l$  au modèle de magnitude obtenu à partir de personnes qui marchent (voir section 3.2) en utilisant la formule (4). Cette probabilité est seuillée pour détecter l'évènement *courir*. un seuil de 0.95 est choisi. Ceci est justifié par le fait qu'il y a une probabilité de 0.95 pour une variable d'être entre  $\mu - 2\sigma$  et  $\mu + 2\sigma$  où  $\mu$  et  $\sigma$  sont respectivement la moyenne et l'écart-type de la loi Gaussienne. Comme  $P_{cour} = 1 - P_{march}$ , l'évènement *marcher* se produit quand il n'y a pas d'évènement *courir* et vice-versa.

#### 4.2 Évènements de rassemblement ou de séparation

Pour la détection des évènements *rassemblement* et *séparation*, on calcule la variance circulaire  $S_{0,f}$  des orientations globales des groupes dans chaque image  $f$  selon l'équation suivante (L.Gaile et E.Burt, 1980) :

$$S_{0,f} = 1 - \frac{1}{n_f} \sum_{i=1}^{n_f} \cos(X_{i,f} - \overline{X_{0,f}}) \quad (7)$$

La variance  $S_{0,f}$  est comprise entre 0 et 1 inclus. Si les angles sont identiques,  $S_{0,f}$  sera égal à 0. Un ensemble d'angles totalement opposés donnera une valeur de  $S_{0,f}$  égale à 1. Si la variance circulaire dépasse un certain seuil  $\beta$  (on choisit  $\beta = 0.3$  dans notre implémentation), on déduit la réalisation d'une *séparation* et/ou d'un *rassemblement*.

On examine aussi la position et l'orientation de chaque groupe par rapport aux autres groupes afin de décider précisément quel évènement s'est produit. Si, par exemple, deux groupes proches sont orientés vers la même destination alors la probabilité de l'évènement *rassemblement* sera grande.

Pour l'évènement *séparation*, on distingue trois situations différentes :

1. La division se produit dans un petit secteur et elle est temporaire, c'est une *dispersion locale*.

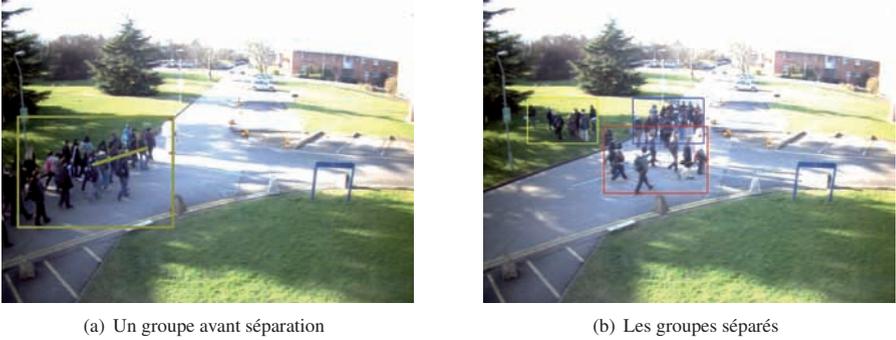


FIG. 6 – Détection d'une séparation

2. La division est plus longue dans le temps et les groupes s'éloignent, c'est une *séparation*.
3. Si la première situation se produit lorsque la foule court, alors c'est une *évacuation*.

Les probabilités de *fusion*, *séparation*, *dispersion locale* et *évacuation* pour l'image  $f$  notés respectivement  $Pfus_f$ ,  $Psep_f$ ,  $Pdisp_f$ ,  $Pevac_f$  sont nuls si la variance circulaire est inférieure à un seuil. Dans le cas contraire, les probabilités de fusion dispersion locale et séparation sont calculés en divisant le nombre d'occurrences de chaque évènement par leur nombre d'occurrences total dans l'image  $f$ . Soient  $Nfus_f$ ,  $Nsep_f$ ,  $Ndisp_f$  le nombre d'occurrences des évènements fusion, séparation, dispersion locale respectivement dans l'image  $f$ . La probabilité de fusion  $Pfus_f$  à l'image  $f$  est calculée comme suit :

$$Pfus_f = \frac{Nfus_f}{Nfus_f + Nsep_f + Ndisp_f} \quad (8)$$

Finalement, la probabilité d'évacuation à l'image  $f$ , notée  $Pevac_f$ , est un cas particulier puisqu'il est conditionné par l'évènement *courir* en plus de l'évènement dispersion locale. Par conséquent, s'il y a un évènement *courir* à l'image  $f$ , alors  $Pdisp_f$  est remplacé par  $Pevac_f$  et  $Ndisp_f$  est remplacé par  $Nevac_f$ .

## 5 Expérimentations

L'approche décrite dans les sections précédentes a été évaluée sur la base de vidéos de PETS'2009. Ces vidéos incluent des séquences contenant différentes activités de foule. Différents scénarios impliquant le calcul de la densité de foule, comptage du nombre de personnes, suivi d'une seule personne dans une foule, analyse de flux et détection d'évènements haut niveau. Un ensemble de données d'apprentissage est aussi disponible, il a été utilisé pour la création des modèles de magnitude et d'orientation de la scène. Nous avons traité les séquences d'analyse de flux et de reconnaissance des évènements. Ces séquences sont au format JPEG avec une résolution de  $720 \times 576$  pixels et une cadence de 8 ips.

## Détection d'évènements dans une scène de foule

Séquence	Description
Séq 1	7 personnes se déplacent en zigzag
Séq 2	Une foule se déplace tout en évitant des obstacles humains
Séq 3	3 des groupes se fusionnent en un seul groupe
Séq 4 et 5	3 personnes circulent dans le sens inverse de la foule

TAB. 1 – Description des séquences utilisées pour la détection des flux multiples

séquence	Marcher	Courir	Dispersion locale	Séparation	Fusion	Évacuation
Séq 6	•	•				
Séq 7	•		•		•	
Séq 8	•			•		
Séq 9	•	•			•	•

TAB. 2 – Description des séquences utilisées pour la détection d'évènements

Les séquences décrites dans le tableau 1 ont été utilisées pour la détection des flux multiples tandis que celles du tableau 2 contiennent les évènements expérimentés. La majorité des évènements ont été détectés avec succès. Quelques fausses détections surviennent, principalement dues à la mesure de profondeur qui donne des distances parfois imprécises entre les groupes et à l'ombre qui génère de faux groupes. Le tableau 3 montre les mesures de précision et de rappel pour les divers évènements expérimentés et qui montre que l'approche donne des résultats satisfaisants par comparaison à la vérité terrain.

	Marcher	Courir	Evacuation	Dispersion locale	Fusion	Séparation
Précision	97,23%	75,86%	69,09%	67,40%	59,14%	47,36%
Rappel	96,17%	81,48%	82,60%	45,04%	45,32%	47,36%

TAB. 3 – Performances du système pour chaque évènement

## 6 Conclusions

Les approches traditionnelles pour l'analyse de comportement se focalisent généralement sur une seule personne. La plupart de ces approches échouent sur les scènes de foule. Dans ce papier, une approche d'analyse globale pour la détection des flux multiples et la détection des évènements liés aux foules a été proposée. Certains évènements sont aussi détectés en analysant la relation spatiotemporelle entre les groupes de personne pour chaque image. Trois catégories d'évènements sont visées : course, rassemblement et séparation de personnes. Les expérimentations sur l'échantillon de vidéos PETS'2009 ont montré que notre méthode est très prometteuse pour des scènes de foules et des scènes complexes. Dans les travaux futurs,

nous appliquerons une méthode de suppression des ombres pour réduire l'effet des ombres. En plus, nous prendrons en compte de l'information de profondeur 2,5D ou 3D pour mesurer plus précisément les distances dans l'espace.

**Remerciements :** Ce travail a été réalisé dans le cadre du projet ANR *CAnADA* (2007-2010, Analyse des comportements Anormaux : Alerte, Détection, Action) et le projet européen Miauce (2006-2009, FP6, Call 5, IST-2005-5-033715).

## Références

- Ali, S. et M. Shah (17-22 June 2007). A Lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. *IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR '07*, 1–6.
- Andrade, E. L., S. Blunsden, et R. B. Fisher (2006a). Hidden Markov models for optical flow analysis in crowds. *18th International Conference on Pattern Recognition. ICPR'06 1*, 460–463.
- Andrade, E. L., S. Blunsden, et R. B. Fisher (2006b). Modelling crowd scenes for event detection. *18th International Conference on Pattern Recognition. ICPR'06*, 175–178.
- Boghossian, B. et S. Velastin (1999). Motion-based machine vision techniques for the management of large crowds. *Proceedings of ICECS '99. The 6th IEEE International Conference on Electronics, Circuits and Systems 2*, 961–964.
- Cupillard, F., A. Avanzi, F. Bremond, et M. Thonnat (2004). Video understanding for metro surveillance. *2004 IEEE International Conference on Networking, Sensing and Control 1*, 186–191.
- Davies, A., J. H. Yin, et S. Velastin (1995). Crowd monitoring using image processing. *Electronics & Communication Engineering Journal 7*(1), 37–47.
- Harris, C. et M. Stephens (1988). A combined corner and edge detector. In *Alvey Vision Conference*, pp. 147–152.
- Hu, M., S. Ali, et M. Shah (2008). Detecting global motion patterns in complex videos. In *ICPR'08 : International Conference on Pattern Recognition*.
- Ihaddadene, N. et C. Djeraba (2008). Real-time crowd motion analysis. *ICPR International Conference on Pattern Recognition, Tampa, Florida, USA*.
- Kaewtrakulpong, P. et R. Bowden (2001). An improved adaptive background mixture model for realtime tracking with shadow detection. *2nd European Workshop on Advanced Video Based Surveillance Systems, AVBS01, VIDEO BASED SURVEILLANCE SYSTEMS*.
- L.Gaile, G. et J. E.Burt (1980). *Directional Statistics*. Concepts and techniques in modern geography ; no. 25. Norwich, England : Geo Abstracts.
- Lin, S.-F., J.-Y. Chen, et H.-X. Chao (2001). Estimation of number of people in crowded scenes using perspective transformation. *IEEE Transactions on Systems, Man and Cybernetics, Part A 31*(6), 645–654.

- Lucas, B. et T. Kanade (1981). An iterative image registration technique with an application to stereo vision. *Proceedings of the International Joint Conference on Artificial Intelligence* (1), 674–679.
- Ma, R., L. Li, W. Huang, et Q. Tian (2004). On pixel count based crowd density estimation for visual surveillance. *IEEE Conference Cybernetics and Intelligent Systems I*, 170–173.
- Marana, A., S. Velastin, L. Costa, et R. Lotufo (1997). Estimation of crowd density using image processing. *IEEE Colloquium Image Processing for Security Applications (Digest No. : 1997/074)*, 11/1–11/8.
- Mehran, R., A. Oyama, et M. Shah (2009). Abnormal crowd behavior detection using social force model. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.
- Rahmalan, H., M. S. Nixon, et J. N. Carter (2006). On crowd density estimation for surveillance. In *International Conference on Crime Detection and Prevention*.
- Shi, J. et C. Tomasi (1994). Good features to track. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 593–600.
- Shyu, M.-L., Z. X. abd Min Chen, et S.-C. Chen (2008). Video semantic event/concept detection using a subspace-based multimedia datamining framework. *IEEE transactions on multimedia ISSN 1520-9210 10*, 252–259.
- Wright, J. et R. Pless (2005). Analysis of persistent motion patterns using the 3d structure tensor. In *WACV-MOTION '05 : Proceedings of the IEEE Workshop on Motion and Video Computing (WACV/MOTION'05) - Volume 2*, Washington, DC, USA, pp. 14–19. IEEE Computer Society.

## Summary

Video-surveillance systems are becoming more and more autonomous in the detection and the reporting of abnormal events. This paper presents a method to detect major flows and events in crowd scenes. These detections are performed using a Direction Model constructed from an online mixture of von Mises distributions applied to the orientation of the optical flow vectors. Major flows are then detected by retrieving the most important orientations from the mixture. Several crowd related events are also detected using a probabilistic model applied to the mean motion magnitude of the optical flow vectors on each frame of the scene. In addition, spatiotemporal relationship analysis of the crowd using the direction model according to the category of the event. The results of processing on a video dataset are presented.