

# Association spatio-temporelle avec données manquantes par minimisation d'énergie

Abir EL ABED<sup>1</sup>, Séverine DUBUISSON<sup>1</sup>, Dominique BÉRÉZIAT<sup>2</sup>

<sup>1</sup>Laboratoire d'Informatique de Paris 6 (LIP6/UPMC)  
104 Avenue du Président Kennedy, 75016 Paris, France

<sup>2</sup>LIP6/UPMC, Clime project/INRIA Rocquencourt  
B.P. 105 78153 Le Chesnay Cedex, France  
abir.elabed@lip6.fr

**Résumé** — Le suivi multi-cible en ligne nécessite la résolution de deux problèmes : l'association de données et l'évaluation en ligne des paramètres du modèle dynamique. Dans cet article, on propose une nouvelle méthode d'association de données, utilisant une énergie à minimiser, définie par des critères géométriques permettant de suivre plusieurs objets, déformables ou non, avec mouvements non linéaires et non constants. L'énergie qu'on propose ne nécessite aucun paramètre ni connaissance *a priori* et ses composantes sont extraites à partir de représentations géométriques (surface et distance) construites avec des mesures et des prédictions. On montre que cette méthode est robuste pour associer correctement une ou plusieurs mesures lorsqu'elles sont équidistantes d'une cible et qu'elle gère les problèmes de données manquantes.

**Abstract** — Online multi-target tracking requires the resolution of two problems: data association and online evaluation of the dynamic model parameters. In this paper, we propose a novel method of data association, using an energy to be minimized, defined by a geometrical criteria to track several objects, deformable or not, with non linear and non constant movements. The energy which we propose does not require neither parameter nor prior knowledges and its components are extracted from geometrical representations (surfaces and distances) built with measurements and predictions. We show that this method is robust to associate correctly one or several measurements when they are equidistant from a target and that it manages the problems of missing data.

## 1 Introduction

Le problème d'association de données [1, 2] est fondamental en suivi temporel multi-cibles. Généralement, l'efficacité des méthodes d'association dépend des observations disponibles et des informations *a priori*. Le manque de données est dû, d'une part, à la méconnaissance *a priori* du modèle de mouvement des cibles, d'autre part, à l'intervalle de temps entre deux observations successives. On rencontre des difficultés d'association lorsque l'on possède peu d'informations sur le modèle de mouvement, et que les objets sont relativement similaires. Le problème se complique si une mesure est équidistante de deux cibles. Les méthodes d'associations actuelles ne permettent pas de faire face à l'ensemble de ces contraintes. Pour y remédier, nous proposons dans cet article une nouvelle méthode d'association qui utilise des critères géométriques pour définir une énergie que l'on cherche ensuite à minimiser. Les principaux avantages de notre approche sont qu'elle n'exige ni paramètres ni connaissances *a priori*, et qu'elle est peu consommatrice en temps de calcul. Cet article est organisé comme suit. Dans la section 2, on expose le modèle mathématique et la présentation géométrique des énergies utilisées pour définir la probabilité d'association mesure-cible. Nous appliquons notre algorithme sur un certain nombre de séquences dans la section 3 pour montrer la robustesse de la méthode proposée. Fina-

lement, on donne une conclusion dans la section 4.

## 2 Modélisation de l'approche : minimisation d'énergie

Nous développons dans cette section un nouvel algorithme d'association de données utilisant uniquement des mesures de positions : la seule information concernant les cibles dont nous disposons correspondent aux prédictions de leurs deux positions antérieures ( $t - 1$  et  $t - 2$ ). Nous parlons ici de prédictions car nous nous positionnons dans un contexte de données manquantes. Nous définissons une énergie  $E$  suivant l'évolution du modèle dynamique de la cible, décrit en termes de déplacements dans l'espace. La scène dynamique est observée par un capteur qui peut fournir exactement une observation à l'instant  $t$ , contenant au moins une mesure, pouvant être associée à une cible ou une fausse alarme. Notre but est d'associer une mesure par cible. On appelle  $(y_1, \dots, y_M)$  le vecteur d'observation contenant  $M$  mesures à un instant particulier. Chaque mesure est définie par une position dans le plan de déplacement. Dans la suite, on désigne par  $A$  la position d'une cible  $k$  et par  $y_j$  une mesure.

On distingue deux modèles dynamiques :

(i) le modèle dynamique initial  $\hat{A}_1(t+1) = f_1(\hat{A}(t)) + B_1$ , avec  $B_1$  un bruit gaussien blanc,  $f_1$  une fonction quelconque et  $\hat{A}(t)$

la prédiction de la position de la cible  $A$  à l'instant  $t$  ;  
(ii) le modèle dynamique mis à jour  $\hat{A}_2(t+1) = f_2(y_j) + B_2$ , avec  $B_2$  un bruit gaussien blanc et  $f_2$  une fonction quelconque dont les paramètres sont mis à jour après l'association de la mesure  $y_j$  à la cible.  
L'énergie globale entre la cible  $k$  et la mesure  $y_j$  est définie par :

$$E(k, y_j) = \frac{1}{\sqrt{3}} \sqrt{\sum_{l=1}^3 \alpha_l^2 (E^l(k, y_j))^2} \quad (1)$$

où  $\alpha_l = \frac{1}{\sum_{k=1}^K E^l(k, y_j)}$  est un facteur de pondération.  
Si le mouvement d'une cible est limité à une translation dans une direction, le problème d'association se réduit au calcul de l'énergie  $E^1$ . Cependant, pour des mouvements complexes tels que des déplacements non affines, des mouvements oscillatoires et des vitesses non constantes,  $E^1$  est insuffisante pour associer correctement les mesures. Pour y remédier, nous incorporons une deuxième énergie  $E^2$  qui indique le degré de similarité entre deux modèles dynamiques. La nécessité d'une troisième énergie  $E^3$ , énergie d'évolution de proximité, sera montrée et clarifiée par des descriptions géométriques, à travers l'étude de quelques cas pour lesquels les deux premières énergies ne suffisent pas à associer correctement les données. Finalement, la mesure  $y_j$  est associée à la cible  $k$  en minimisant l'énergie globale :

$$\begin{aligned} \mathcal{D}_{y_j \rightarrow k} &= \left\{ \underset{k=1, \dots, K}{\operatorname{argmin}} E(k, y_j) \right\} \\ &= \left\{ \underset{k=1, \dots, K}{\operatorname{argmin}} \left( \frac{1}{\sqrt{3}} \sqrt{\sum_{l=1}^3 \alpha_l^2 (E^l(k, y_j))^2} \right) \right\} \end{aligned} \quad (2)$$

avec  $0 \leq \alpha_l \leq 1$  et  $0 \leq E(k, y_j) \leq 1$ .

Pour prédire une position, on utilise souvent un modèle dynamique, dont les paramètres sont généralement fixés par apprentissage [3, 4], qui représente ainsi des mouvements réguliers. Pour les mouvements complexes (vitesses non constantes ou oscillations non périodiques), le choix des paramètres est plus difficile, particulièrement quand on a des données manquantes. Pour ces raisons, les paramètres de notre modèle dynamique sont ajustés d'une manière adaptative à chaque fois qu'une mesure est disponible [5]. L'énergie  $E(k, y_j)$  est une combinaison linéaire de trois énergies,  $\{E^1, E^2, E^3\}$ , données par :

1. L'énergie  $E^1(k, y_j)$ , donne la distance de Mahalanobis entre une mesure  $y_j$  disponible à l'instant  $t$  et la prédiction de la position de la cible  $A$  à  $t-1$ . Cette énergie est suffisante si les déplacements de la cible sont rectilignes. Elle est définie par :

$$E^1(k, y_j) = \sqrt{(y_j - \hat{A}(t-1))^T \Sigma_k^{-1} (y_j - \hat{A}(t-1))}$$

où  $\Sigma_k$  est la matrice de covariance de la prédiction de la cible  $k$ , désignée par  $\hat{A}$  dans l'équation. On suppose que les positions sont indépendantes et on fixe leurs variances. On suppose aussi que la matrice de covariance est connue *a priori*.

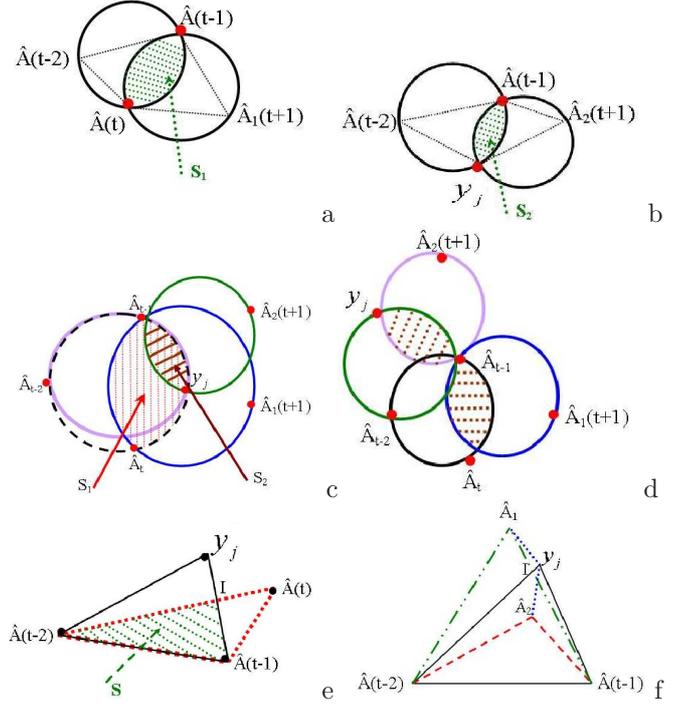


FIG. 1 – (a-b-e) Surfaces d'intersections  $\{S_1, S_2, S\}$ ; (c-d) Différence entre les surfaces  $S_1$  et  $S_2$  extraites à partir de deux modèles dynamiques différents; (g) Surfaces d'intersections dans le cas où deux prédictions sont équidistantes de  $y_j$ .

2. Dans le cas de mouvements complexes, nous ajoutons l'énergie d'évolution absolue  $E^2(k, y_j)$ . Elle quantifie l'exactitude géométrique entre les paramètres de deux modèles dynamiques de prédiction à  $t+1$  dont l'évolution est différente, donnés par :

- (i) le modèle dynamique non mis à jour ( $\hat{A}_1(t+1)$ ) qui ne considère pas la présence de la mesure  $y_j$  ;
- (ii) le modèle dynamique mis à jour ( $\hat{A}_2(t+1)$ ), dont les paramètres sont recalculés en associant à  $t$  la mesure  $y_j$  à la cible  $k$ .

$E^2(k, y_j)$  permet d'évaluer numériquement la proximité entre deux modèles dynamiques. Notre idée repose sur l'évaluation des paramètres du modèle dynamique dans deux cas selon que la mesure  $y_j$  est produite par la cible ou non. Nous prédisons tout d'abord les états,  $\hat{A}_1(t+1)$  et  $\hat{A}_2(t+1)$ , à  $t+1$ . Ensuite, nous déterminons la surface  $S_1$  définie par l'intersection des deux cercles circonscrits aux triangles ( $\hat{A}(t-2), \hat{A}(t-1), \hat{A}(t)$ ) et ( $\hat{A}(t-1), \hat{A}(t), \hat{A}_1(t+1)$ ), et la surface d'intersection  $S_2$  entre les deux cercles circonscrits aux triangles ( $\hat{A}(t-2), \hat{A}(t-1), y_j$ ) et ( $\hat{A}(t-1), y_j, \hat{A}_2(t+1)$ ) (figures 1.(a-b)).

L'énergie  $E^2(k, y_j) = |S_1 - S_2|$  est minimale quand les deux modèles sont proches. Cependant, l'énergie  $E^2$  est-elle capable de gérer tout les types de mouvement ? En regardant son expression, différence de deux surfaces, on se rend compte qu'elle est fiable si  $\hat{A}(t)$  et  $y_j$  se situent du même côté de l'axe ( $\hat{A}_{t-2}\hat{A}_{t-1}$ ) (figure 1.c). Lorsque  $\hat{A}(t)$  et  $y_j$  sont diamétralement opposés ou lorsqu'ils se trouvent de part et d'autre de l'axe ( $\hat{A}_{t-2}, \hat{A}_{t-1}$ ), les surfaces  $S_1$  et  $S_2$  sont sem-

blables (voir figure 1.d), ce qui implique que  $E^2$  est nulle. Dans ces cas, cette énergie n'est pas une source d'information suffisante pour l'association, et nous proposons d'ajouter une troisième énergie  $E^3$ .

3. L'énergie d'évolution de proximité,  $E^3(k, y_j) = \frac{1}{S}$ , est l'inverse de la surface d'intersection des deux triangles  $(\hat{A}(t-2), \hat{A}(t-1), y_j)$  et  $(\hat{A}(t-2), \hat{A}(t-1), \hat{A}(t))$  (surface hachée dans la figure 1.e). Cette énergie évalue la similarité entre la prédiction  $\hat{A}(t)$  et la mesure  $y_j$ . L'accroissement de  $S$  signifie que la prédiction et la mesure à l'instant  $t$  sont proches.

La raison pour laquelle nous travaillons sur une surface d'intersection au lieu de calculer la distance entre la mesure  $y_j$  et la prédiction de la position de la cible à l'instant  $t$  est illustrée par la figure 1.f, sur laquelle les deux prédictions à l'instant  $t$ ,  $\hat{A}_1$  et  $\hat{A}_2$ , sont équidistantes de la mesure  $y_j$ . Si nous calculons la distance pour mesurer l'énergie de proximité, nous obtenons que les deux modèles ont le même degré de similarité avec le modèle initial défini par le modèle dynamique des points  $(\hat{A}(t-2), \hat{A}(t-1), y_j)$ . Ce résultat mène à une contradiction avec la réalité. Ce problème peut être expliqué dans cet exemple par le fait que si elles ont toutes les deux le même degré de similarité avec le troisième modèle dynamique, nous pouvons conclure que leurs cibles correspondantes ont la même dynamique. Pour cette raison, nous avons choisi également d'évaluer la similarité en extrayant la surface d'intersection entre les triangles qui sont très différentes (voir figure 1.f), ce qui mène à une association correcte.

Nous avons décrit une approche par minimisation d'énergie permettant d'associer les mesures d'une observation à des cibles. Les composantes de cette énergie sont extraites à partir des représentations géométriques (surfaces et distances) construites avec des mesures et des prédictions. Ce choix de représentation géométrique pour définir ces énergies est motivé par les volontés de :

- (i) montrer la continuité géométrique du système entre les prédictions et les états précédents en utilisant deux modèles dynamiques différents ;
- (ii) mesurer la similarité entre les prédictions, à un moment particulier, en utilisant deux modèles dynamiques différents, qui logiquement doivent être quasiment semblables parce qu'ils représentent le même système.

### 3 Résultats expérimentaux

La figure 2 montre les acquisitions aux instants  $t-2, t-1$  et  $t$ , respectivement, et contient deux cibles  $T_1$  et  $T_2$ . Une observation est disponible à l'instant  $t$ , contenant deux mesures  $M_1$  et  $M_2$  (positions dans l'espace  $(x, y)$ ). Pour associer ces deux mesures à leur cible, nous calculons dans un premier temps la distance de Mahalanobis et obtenons  $\{\alpha_1 E^1(T_1, M_2) = 0.32\} < \{\alpha_1 E^1(T_2, M_2) = 0.68\}$ , signifiant que la distance de  $M_2$  à  $T_1$  est plus petite que celle de  $M_2$  à  $T_2$  (ce qui est une contradiction avec la réalité).

En calculant les autres énergies  $E^2$  et  $E^3$ , on obtient :

$$E(T_1, M_1) = \frac{1}{\sqrt{3}} \sqrt{(0.31)^2 + (0.02)^2 + (0.34)^2} = \underline{0.27}$$

$$E(T_1, M_2) = \frac{1}{\sqrt{3}} \sqrt{(0.32)^2 + (0.72)^2 + (0.86)^2} = 0.68$$

$$E(T_2, M_1) = \frac{1}{\sqrt{3}} \sqrt{(0.69)^2 + (0.98)^2 + (0.66)^2} = 0.79$$

$$E(T_2, M_2) = \frac{1}{\sqrt{3}} \sqrt{(0.68)^2 + (0.28)^2 + (0.14)^2} = \underline{0.43}$$

D'après la première énergie de Mahalanobis, on a :

$$\begin{aligned} \{\alpha_1 E^1(T_1, M_1) = 0.31\} &< \{\alpha_1 E^1(T_2, M_1) = 0.69\} \\ \{\alpha_1 E^1(T_1, M_2) = 0.32\} &< \{\alpha_1 E^1(T_2, M_2) = 0.68\} \end{aligned}$$

Ce qui désigne que les deux mesures sont plus proche de la cible  $T_1$  et dans ce cas les deux mesures seront associées à cette cible si on utilise la méthode d'association par minimisation de distance. Ce résultat produit une contradiction avec la réalité. En utilisant notre approche, on remarque que  $E(T_1, M_1) < E(T_2, M_1)$  et  $E(T_2, M_2) < E(T_1, M_2)$ , ce qui se traduit par  $M_1$  et  $M_2$  sont respectivement associées à  $T_1$  et à  $T_2$ . On rappelle qu'une mesure  $M_j$  est associée à la cible  $T_i$  lorsque l'amplitude de l'énergie globale  $E(T_i, M_j)$  est minimale.



FIG. 2 – Acquisitions aux instants  $\{t-2, t-1, t\}$ .

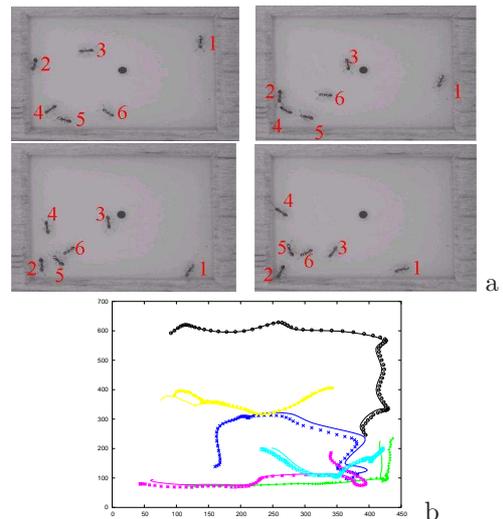


FIG. 3 – (a) Acquisitions à  $\{t-2, t-1, t$  et  $t+1\}$  (6 cibles) et résultats d'association ; (b) Acquisitions à  $\{t-2, t-1, t\}$  (2 cibles) ; (b) Estimation des trajectoires des fourmis.

La seconde séquence testée montre des fourmis dont leur déplacement est erratique avec une vitesse non constante (figure 3.b) : elles changent de direction, accélèrent, ralentissent, s'arrêtent, font des rotations autour leur axe, etc.

L'observation disponible à l'instant  $t$  contient six mesures (positions dans l'espace  $(x, y)$ ). Du fait de la similarité des cibles, seul leur mouvement peut aider à les associer aux mesures disponibles. La figure 3.a montre les acquisitions obtenues aux instants  $\{t-2, t-1, t, t+1\}$  respectivement et qui correspondent aux images  $\{10, 25, 35, 45\}$  de la séquence. Les fourmis sont numérotées pour montrer leurs déplacements d'une acquisition à une autre. Le tableau 1 donne les valeurs numériques des composantes d'énergie,

$\begin{pmatrix} \alpha_1 E^1(T_i, M_j) \\ \alpha_2 E^2(T_i, M_j) \\ \alpha_3 E^3(T_i, M_j) \end{pmatrix}$ , quand une mesure  $M_j$  est associée à une cible  $T_i$ . Si nous utilisons la méthode d'association basée uniquement sur la minimisation de distance, on obtient les résultats suivants :

$$\alpha_1 E^1(T_1, M_1) = \operatorname{argmin}_{i=1, \dots, 6} \{\alpha_1 E^1(T_i, M_1)\} \quad (3)$$

$$\alpha_1 E^1(T_4, M_2) = \operatorname{argmin}_{i=1, \dots, 6} \{\alpha_1 E^1(T_i, M_2)\} \quad (4)$$

$$\alpha_1 E^1(T_3, M_3) = \operatorname{argmin}_{i=1, \dots, 6} \{\alpha_1 E^1(T_i, M_3)\} \quad (5)$$

$$\alpha_1 E^1(T_2, M_4) = \operatorname{argmin}_{i=1, \dots, 6} \{\alpha_1 E^1(T_i, M_4)\} \quad (6)$$

$$\alpha_1 E^1(T_2, M_5) = \operatorname{argmin}_{i=1, \dots, 6} \{\alpha_1 E^1(T_i, M_5)\} \quad (7)$$

$$\alpha_1 E^1(T_6, M_6) = \operatorname{argmin}_{i=1, \dots, 6} \{\alpha_1 E^1(T_i, M_6)\} \quad (8)$$

On remarque que les mesures  $\{M_1, M_3, M_6\}$  sont correctement associées à leur cible  $\{T_1, T_3, T_6\}$ . Parcontre, les mesures  $\{M_2, M_4, M_5\}$  seraient respectivement associées à  $\{T_4, T_2, T_2\}$  ce qui est faux (voir les équations (4,6 et 7)). Pour compenser ces résultats et pour mieux associer ces mesures, on calcul les deux autres énergies  $\alpha_2 E^2$  et  $\alpha_3 E^3$ , on obtient  $\alpha_2 E^2(T_2, M_2) < \alpha_2 E^2(T_4, M_2)$  et  $\alpha_3 E^3(T_2, M_2) < \alpha_3 E^3(T_4, M_2)$  ce qui compense l'erreur donnée par  $\alpha_1 E^1(T_2, M_2)$ . Finalement, l'amplitude de l'énergie globale  $E(T_i, M_2)$  est minimale quand  $M_2$  est associée à  $T_2$ . Cette séquence nous montre la nécessité de l'énergie  $E^3$  dans notre formulation. A titre d'exemple, si l'énergie globale est composée seulement des deux énergies  $\alpha_1 E^1(T_i, M_j)$  et  $\alpha_2 E^2(T_i, M_j)$  pour associer les mesures, l'énergie globale dans ce cas est donnée par :

$$E(T_i, M_j) = \frac{1}{\sqrt{2}} \sqrt{(\alpha_1 E^1(T_i, M_j))^2 + (\alpha_2 E^2(T_i, M_j))^2}.$$

Si on se limite à l'utilisation des deux énergies, on obtient :  $E(T_6, M_5) = 9,22$  et  $E(T_5, M_5) = 8,2$ . D'où  $E(T_6, M_5) < E(T_5, M_5)$  et la mesure  $M_5$  serait associée à la cible  $T_6$ , ce qui provoque une mauvaise association. Avec l'intégration de la troisième énergie, on obtient  $E(T_5, M_5) < E(T_6, M_5)$  ce qui compense les lacunes des autres énergies. Finalement, dans le tableau 2, nous observons que chaque mesure est correctement associée à sa cible. Le temps de calcul de toutes ces énergies est de 0.25 secondes sous **Matlab**.

## 4 Conclusion

Nous avons proposé une nouvelle méthode d'association de données basée sur la minimisation d'une énergie globale mesurant le degré de similarité de deux modèles dy-

namiques. L'approche développée est robuste aux mouvements complexes et aux systèmes non linéaires avec données manquantes. Son principal atout est qu'elle n'exige aucun paramètre. Comme perspectives, nous suggérons d'intégrer l'approche de minimisation d'énergie dans le filtre particulière pour le suivi multicible et d'ajouter des composantes d'énergie permettant de traiter des observations multimodales. Cette phase est actuellement en cours de développement.

	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	$M_6$
$T_1$	6.5	47.2	25.8	43.9	48.7	46.6
	1.5	1.8	1.4	1.1	11.3	1.1
	0.03	45.8	0.8	14.1	48.1	43.2
$T_2$	22.5	5.8	21.4	<u>6.5</u>	<u>2.5</u>	12.1
	3.1	1.2	3.1	9.6	54.5	1.6
	6.8	0.01	24.6	0.2	14.6	13.6
$T_3$	15.1	24.9	4.1	17.7	23.7	19.7
	14.2	0.2	0.3	0.2	1.3	0.1
	22.2	8.5	1.8	5.6	4.7	24.5
$T_4$	21.4	<u>1.7</u>	21.6	9.4	5.4	10.2
	6.6	2.2	2.7	0.3	24.1	4.4
	38.8	14.8	27.8	0.7	17.8	12.1
$T_5$	18.3	6.1	17.9	12.2	<u>9.3</u>	5.9
	74.4	85.8	92.2	96.2	<u>6.9</u>	92.4
	8.5	10.5	12.6	3.2	<u>4.05</u>	6.2
$T_6$	16.1	13.5	9.3	10.9	<u>10.3</u>	5.5
	0.3	0.25	0.2	1.1	<u>1.9</u>	0.4
	23.7	20.1	32.4	76.3	<u>10.8</u>	0.35

TAB. 1 – Valeur numérique des différentes énergies quand une mesure  $M_j$  est associée à une cible  $T_i$ .

	$E(T_i, M_j)$					
$T_1$	<u>3.9</u>	38.1	14.9	26.7	40.1	36.7
$T_2$	13.7	<u>3.4</u>	18.9	6.7	32.6	10.5
$T_3$	17.5	15.2	<u>2.6</u>	10.8	14	18.1
$T_4$	25.9	8.7	20.4	<u>5.4</u>	17.6	9.5
$T_5$	44.4	50.1	54.7	56.1	<u>7.1</u>	53.6
$T_6$	16.6	13.9	19.4	44.5	8.7	<u>3.2</u>

TAB. 2 – L'amplitude des énergies globales quand une mesure  $M_j$  est associée à une cible  $T_i$ .

## Références

- [1] Rong, L., Bar-Shalom, Y. : Tracking in clutter with nearest neighbor filter : analysis and performance. In : IEEE transactions on aerospace and electronic systems. (1996)
- [2] Vermaak, J., Godsill, S., Pérez, P. : Monte carlo filtering for multi-target tracking and data association. In : IEEE Transactions on Aerospace and Electronic Systems. (2005)
- [3] North, B., Blake, A., Isard, M., Rittscher, J. : Learning and classification of complex dynamics. In : IEEE Transactions on Pattern Analysis and Machine Intelligence. (2000)
- [4] Blake, A., Isard, M. : Active contours. In : Springer-Verlag. (1998)
- [5] Abed, A.E., Dubuisson, S., Béréziat, D. : Comparison of statistical and shape-based approaches for non-rigid motion tracking with missing data using a particle filter. In : Advanced Concepts for Intelligent Vision Systems. (2006)