



Extraction de pistes multiples dans le contexte du sonar passif

H. GAUVRIT*, C. JAUFFRET** et J.P LE CADRE*

* IRISA/CNRS, Campus de Beaulieu, 35042 Rennes Cedex, France

** DCN/Ingénierie/Sud, Le Brusac, 83140, Six-Fours-Les-Plages, France

RESUME

Dans les approches classiques du problème d'extraction de pistes multiples, on est confronté au problème d'explosion combinatoire des hypothèses d'association des mesures aux pistes. On montre que l'algorithme PMHT remédie à cette explosion en modélisant ces associations par des variables aléatoires qu'il s'agit d'estimer conjointement aux états des sources au moyen de l'algorithme EM (Expectation-Maximisation). On applique ensuite cette approche au cas de modèles non-linéaires des observations.

ABSTRACT

Traditional multi-hypothesis tracking algorithms need an enumeration of all the assignments of measurements to tracks which lead to exponential computational complexity. It is shown that the PMHT algorithm avoid this enumeration by estimating the probabilities that the measurements belong to tracks via the Expectation-Maximisation (EM) algorithm. This method is then applied to non-linear models of measurements

1 Introduction

Dans une chaîne de traitement de l'information d'un système radar ou sonar, l'extraction intervient à l'issue des étapes de traitement du signal (analyse spectrale, traitement d'antenne, etc...). Elle a pour but d'élaborer des pistes à partir des observations issues du traitement des signaux en sortie de capteurs. Ces pistes servent ensuite à la localisation des sources, à la classification, voire à l'optimisation de la tactique... La difficulté de cette étape réside essentiellement dans l'incertitude sur l'origine même des observations. En effet, les mesures dont on dispose lors de cette étape sont de deux types: soit ce sont des fausses-alarms, soit elles proviennent des sources que l'on cherche à caractériser.

L'abondante littérature relative à l'extraction a essentiellement pour objet le domaine radar. Les algorithmes proposés jusqu'à présent (JPDAF, MHT) utilisent une énumération exhaustive de toutes les assignations possibles des mesures aux pistes. On est alors confronté à un problème d'explosion combinatoire qui se résoud par élimination des séquences les moins probables (principe de fenêtrage dans le cas du JPDAF et d'élagage pour le MHT). On se propose dans cet article de présenter une nouvelle approche initialement proposée par R. Streit [6] et de l'appliquer au cas de modèles non-linéaires des observations. Cette méthode consiste à modéliser les assignations comme des variables aléatoires. Les mesures ne sont plus associées à des pistes précises comme c'est le cas dans les méthodes usuelles mais elles sont, au contraire, estimées au moyen de l'algorithme EM. On désigne ce type de méthode sous le vocable général PMHT (Probabilistic Multi-Hypothesis Tracking). Après avoir formulé le problème d'extraction et exprimé les hypothèses contenues dans le PMHT, on rappelle le principe de l'algorithme EM que l'on applique ensuite à l'extraction de pistes multiples.

2 Formulation du problème

On considère que l'on dispose de M modèles de sources manœuvrantes¹. L'algorithme PMHT est un algorithme batch, soit T sa durée.

A chaque instant $t = 1 \dots T$, on doit estimer M états, notés $X(t) = (x_1(t), \dots, x_M(t))$. L'ensemble des mesures observées à l'instant t , appelé scan, est noté $Z(t) = (z_1(t), \dots, z_{m_t}(t))$ avec m_t , le nombre de mesures observées à l'instant t .

L'ensemble des états estimés et l'ensemble des mesures collectées tout au long du batch sont, quant à eux, notés:

$$\mathcal{X} = (X(1), \dots, X(T))$$

$$\mathcal{Z} = (Z(1), \dots, Z(T))$$

On introduit de plus, un vecteur supplémentaire, contenant l'information sur les assignations des mesures aux pistes, on le désignera par $\mathcal{K} = (K(1), \dots, K(T))$ pour spécifier les associations définies sur tout le batch, et chaque vecteur d'assignation à un instant t par $K(t) = (k_1(t), \dots, k_{m_t}(t))$ où $k_j(t) = m$ signifie que la $j^{\text{ème}}$ mesure du scan à l'instant t est associé au $m^{\text{ème}}$ modèle source.

Le vecteur de paramètres que l'on va chercher à estimer est noté \mathcal{O} . Il contient bien entendu, le vecteur des états des modèles, \mathcal{X} , mais aussi les probabilités d'assignation d'une mesure à un modèle que l'on notera, Π . Il faut bien remarquer que l'idée originale [6] de cette méthode est d'inclure au vecteur d'état, le vecteur des probabilités d'assignation. Donc, $\mathcal{O} = (\mathcal{X}, \Pi)$ et en respectant la même notation que celle définie jusqu'alors, on désigne le vecteur de paramètres à un instant donné par: $O(t) = (X(t), \Pi(t))$, où $\Pi(t) = (\pi_1(t), \dots, \pi_M(t))$. La notation $\pi_m(t)$ désigne la probabilité d'associer une mesure au modèle m . Cela signifie que cette probabilité est indépendante de la mesure, ie: $\pi_m(t) = p(k_j(t)=m), \forall j = 1 \dots m_t$.

3 Les hypothèses du PMHT

On considère que les scans observés à deux instants distincts et que les mesures à l'intérieur même d'un scan sont

1. Un des M modèles peut modéliser les fausses-alarms.



indépendants conditionnellement à \mathcal{O} . La première partie de la proposition est une hypothèse classique en extraction ce qui n'est pas le cas de la seconde partie de l'énoncé. En effet, dans les travaux de [1], [3], [4], on considère

- qu'une mesure est associée à une piste ou à une fausse-alarme ce qui traduit l'exclusivité et l'exhaustivité de l'association d'une mesure.
- qu'une piste est à l'origine d'au plus une mesure ce qui met en évidence la dépendance des mesures à l'intérieur d'un scan.

Seule la première hypothèse est retenue dans le cas du PMHT. Avec la notation du paragraphe précédent, cette condition amène la contrainte suivante sur les probabilités d'assignation:

$$\sum_{m=1}^M \pi_m(t) = 1.$$

Puisqu'aucune contrainte n'est appliquée sur l'assignation des mesures d'un scan aux pistes, toutes les mesures peuvent très bien être associées à une unique source. Cette hypothèse retenue dans le cas du PMHT est originale et tout à fait réaliste puisqu'une source peut être à l'origine de plusieurs mesures. On prend donc en compte beaucoup plus d'hypothèses d'associations mesures-pistes que dans les approches classiques type MHT, JPDAF...

A l'intérieur d'un scan à un instant t , le vecteur de mesure $Z(t)$ et le vecteur d'assignation $K(t)$ sont indépendants. On suppose aussi que le vecteur $K(t)$ est indépendant des états des sources, $X(t)$, à cet instant. De plus, les vecteurs d'états des différents modèles sont indépendants entre eux. Cette dernière hypothèse nous sera utile pour décomposer la fonction à maximiser suivant chacun des modèles.

4 L'algorithme EM

4.1 Généralités

On rappelle le principe de l'algorithme présenté par Dempster, Laird et Rubin dans [2]. Soient deux espaces \mathcal{X} et \mathcal{Y} et une application de \mathcal{X} dans \mathcal{Y} . On notera $\mathcal{X}(y) = \{x \mid x \in \mathcal{X} \text{ et } y(x) = y\}$. On suppose qu'on ne peut pas accéder directement à la mesure de x . Cela correspond à de nombreux problèmes physiques où des données sont manquantes par exemple. On qualifie x de "données complètes", et y de "données incomplètes". On note, $g(y|\phi)$ la densité de y et $f(x|\phi)$ la densité de x et $\phi \in \Omega$ le vecteur de paramètres inconnus à estimer par le maximum de vraisemblance. Alors,

$$g(y|\phi) = \int_{\mathcal{X}(y)} f(x|\phi) dx.$$

L'algorithme EM est une méthode pour déterminer une valeur de ϕ qui maximise $g(y|\phi)$ pour une observation y . Chaque itération de l'algorithme comprend une étape E (Expectation) et une étape M (Maximization).

1. Lors de l'étape E, à partir des paramètres ϕ^i de l'itération précédente, on calcule

$$Q(\phi|\phi^i) = E\{\log(f(x|\phi))\mid y, \phi^i\}$$

2. L'étape M maximise cette expression de façon à mettre à jour les nouveaux paramètres pour l'étape suivante soit

$$\phi^{i+1} = \arg \max_{\phi} Q(\phi|\phi^i)$$

On peut interpréter l'algorithme de la façon suivante: ne connaissant pas la fonction de vraisemblance des données complètes, on l'estime à partir des observations, y , et des paramètres estimés disponibles à l'étape courante. On peut montrer qu'en combinant ces deux étapes la fonction log-vraisemblance, notée $L(\phi) = \log(g(y|\phi))$, croît [2]. Une étude des critères de convergence est présentée dans [7].

4.2 Un exemple d'application: le cas de mélanges de densités

On s'intéresse dans ce paragraphe à l'application de l'algorithme EM à un problème particulier, celui des mélanges de densités [5]. On verra au paragraphe suivant que le problème qui nous intéresse peut être vu comme un problème de mixture.

x désigne toujours les "mesures complètes" et y les "mesures incomplètes". On suppose avoir observé N mesures indépendantes les unes des autres, $\{y_j\}_{j=1..N}$. Chacune de ces mesures appartient à une famille paramétrée de densités de probabilité de la forme

$$p(y_j|\phi) = \sum_{i=1}^m \pi_i p_i(y_j|\phi_i)$$

où chaque π_i est non-négatif et vérifie $\sum_{i=1}^m \pi_i = 1$ et où chaque p_i est une densité de probabilité paramétrée par ϕ_i . On note $\phi = (\pi_1, \dots, \pi_m, \phi_1, \dots, \phi_m)$. Cela signifie que chaque mesure, y_j , provient d'une des m densités $\{p_i\}_{i=1..m}$ avec pour probabilité $\{\pi_i\}_{i=1..m}$.

La fonction de vraisemblance des mesures incomplètes et la fonction log-vraisemblance s'écrivent alors respectivement:

$$g(y|\phi) = \prod_{j=1}^N \sum_{i=1}^m \pi_i p_i(y_j|\phi_i)$$

$$L(\phi) = \sum_{j=1}^N \log\left(\sum_{i=1}^m \pi_i p_i(y_j|\phi_i)\right)$$

Etape E

Le vecteur des données complètes peut être défini pour chaque mesure j , par $x_j = (y_j, k_j)$ où k_j est un indice compris entre 1 et m indiquant l'origine de la mesure y_j . Par conséquent,

$$f(x|\phi) = p(y|k, \phi) p(k|\phi)$$

$$= \prod_{j=1}^N \pi_{k_j} p_{k_j}(y_j|\phi_{k_j})$$

A l'itération $(i+1)$, il s'agit de calculer à l'étape E, l'expression $Q(\phi|\phi^i) = E\{\log(f(x|\phi))\mid y, \phi^i\}$:

$$Q(\phi|\phi^i) = \sum_k \log(f(x|\phi)) p(k\mid y, \phi^i)$$

$$= \sum_{k_1=1}^m \dots \sum_{k_N=1}^m \left\{ \sum_{j=1}^N \log[\pi_{k_j} p_{k_j}(y_j|\phi_{k_j})] \right\}$$

$$\prod_{j'=1}^N \frac{\pi_{k_{j'}}^i p_{k_{j'}}(y_{j'}|\phi_{k_{j'}}^i)}{p(y_{j'}|\phi^i)} \quad (1)$$

En considérant les N sommations sur un des éléments de l'expression entre accolades et en regroupant judicieusement les termes, l'expression se simplifie sous la forme:

$$Q(\phi|\phi^i) = \sum_{k=1}^m \left[\sum_{j=1}^N \frac{\pi_k^i p_k(y_j|\phi_k^i)}{p(y_j|\phi^i)} \right] \log \pi_k + \sum_{k=1}^m \sum_{j=1}^N \log p_k(y_j|\phi_k) \frac{\pi_k^i p_k(y_j|\phi_k^i)}{p(y_j|\phi^i)} \quad (2)$$

Remarques: L'expression (2) se maximise suivant π et (ϕ_1, \dots, ϕ_m) indépendamment. Il faut noter de plus, qu'à la condition que les paramètres ϕ_i sont mutuellement indépendants, la maximisation suivant (ϕ_1, \dots, ϕ_m) se décompose de nouveau en m maximisations, chacune d'elles étant associée à un des paramètres ϕ_i . Enfin, puisque $\log \pi_1, \dots, \log \pi_m$ apparaissent linéairement dans (2), on peut déterminer une solution unique au premier problème de maximisation.

Etape M

1. La maximisation suivant π de l'expression

$$g(\pi) = \sum_{k=1}^m \left[\sum_{j=1}^N \frac{\pi_k^i p_k(y_j|\phi_k^i)}{p(y_j|\phi^i)} \right] \log \pi_k$$

sous la contrainte:

$$\sum_{k=1}^m \pi_k = 1.$$

amène la nouvelle probabilité d'assignation remise à jour:

$$\pi_k^{i+1} = \frac{1}{N} \sum_{j=1}^N \frac{\pi_k^i p_k(y_j|\phi_k^i)}{p(y_j|\phi^i)} \quad (3)$$

2. Dans un second temps, on maximise, suivant $\{\phi_k\}_{k=1..m}$, l'expression

$$\sum_{k=1}^m \sum_{j=1}^N \log p_k(y_j|\phi_k) \frac{\pi_k^i p_k(y_j|\phi_k^i)}{p(y_j|\phi^i)}$$

Le nouvel état mis à jour s'exprime sous la forme:
 $\forall k = 1..m,$

$$\phi_k^{i+1} \in \text{Arg max}_{\phi_k} \sum_{j=1}^N \log p_k(y_j|\phi_k) \frac{\pi_k^i p_k(y_j|\phi_k^i)}{p(y_j|\phi^i)} \quad (4)$$

Chaque poids $\frac{\pi_k^i p_k(y_j|\phi_k^i)}{p(y_j|\phi^i)}$ correspond à la probabilité a posteriori que la mesure y_j provienne de la k^{eme} population sachant l'estimation courante du maximum de vraisemblance ϕ^i . Le raisonnement mené est étendu au problème d'extraction de pistes multiples dans la section suivante.

5 L'algorithme PMHT

Sur toute la longueur du batch de durée T , on observe les mesures \mathcal{Z} . Mais, on ne peut pas accéder à la connaissance du vecteur des assignations \mathcal{K} . On qualifie donc le couple $(\mathcal{Z}, \mathcal{K})$ de mesures "complètes" et \mathcal{Z} de mesures "incomplètes". La fonction de vraisemblance des observations "complètes" s'écrit alors:

$$\begin{aligned} p(\mathcal{Z}, \mathcal{K}|\mathcal{O}) &= \prod_{t=1}^T p(\mathcal{Z}(t), \mathcal{K}(t)|\mathcal{O}(t)) \\ &= \prod_{t=1}^T \prod_{j=1}^{m_t} p(z_j(t)|x_{k_j}(t)) \pi_{k_j}(t) \end{aligned} \quad (5)$$

et la vraisemblance des observations "incomplètes":

$$\begin{aligned} p(\mathcal{Z}|\mathcal{O}) &= \prod_{t=1}^T p(\mathcal{Z}(t)|X(t), \Pi(t)) \\ &= \prod_{t=1}^T \prod_{j=1}^{m_t} p(z_j(t)|\mathcal{O}(t)) \\ &= \prod_{t=1}^T \prod_{j=1}^{m_t} \sum_{m=1}^M p(z_j(t)|x_m(t)) \pi_m(t) \end{aligned} \quad (6)$$

L'expression (6) provient de l'hypothèse particulière du PMHT d'indépendance des mesures à l'intérieur d'un scan. L'algorithme consiste à estimer non seulement les états des sources mais aussi les probabilités d'association. On rappelle que ces dernières sont soumises à la contrainte

$$\sum_{m=1}^M \pi_m(t) = 1.$$

5.1 Etape E

L'étape E de l'algorithme EM consiste à calculer

$$\begin{aligned} Q(\mathcal{O}|\mathcal{O}^*) &= E(\log p(\mathcal{Z}, \mathcal{K}|\mathcal{O}) | \mathcal{Z}, \mathcal{O}^*) \\ &= \sum_{\mathcal{K}} \log p(\mathcal{Z}, \mathcal{K}|\mathcal{O}) p(\mathcal{K}|\mathcal{Z}, \mathcal{O}^*) \end{aligned} \quad (7)$$

où \mathcal{O}^* désigne le vecteur de paramètres estimé lors de l'étape précédente. Ce calcul est une extension de celui mené dans le cas de mixture de densités où la collection de mesures obtenues à un instant donné est généralisée sur toute la longueur du batch de durée T . En définissant le poids de la j^{eme} mesure associée au modèle m , $w_{j,k_j}(t)$

$$w_{j,k_j}(t) = p(k_j(t)|z_j(t), \mathcal{O}^*(t)) = \frac{\pi_{k_j}^* p(z_j(t)|x_{k_j}^*(t))}{p(z_j(t)|\mathcal{O}^*(t))}$$

On obtient

$$\begin{aligned} Q(\mathcal{O}|\mathcal{O}^*) &= \sum_{k=1}^M \sum_{t=1}^T \left[\sum_{j=1}^{m_t} w_{j,k}(t) \right] \log \pi_k(t) \\ &+ \sum_{k=1}^M \sum_{t=1}^T \sum_{j=1}^{m_t} \log [p(z_j(t)|x_k(t))] w_{j,k}(t) \end{aligned} \quad (8)$$

A partir de l'expression (8), plusieurs remarques s'imposent:

- L'étape de maximisation, M, se divise en deux maximisations: l'une suivant Π et l'autre suivant \mathcal{X} .



960 De plus, la maximisation peut être conduite séparément pour chaque modèle. Il faut cependant noter que les maximisations sont couplées au moyen de l'expression $w_{j,k}(t) = \frac{\pi_k^* p(z_j(t)|x_k^*(t))}{p(z_j(t)|O^*(t))}$.

5.2 Etape M

1. La première maximisation, suivant Π , est identique à celle du paragraphe précédent. La seule différence, c'est que les probabilités d'association varient au cours du temps.

$$\forall t = 1 \dots T, \pi_k(t) = \frac{1}{m_t} \sum_{j=1}^{m_t} w_{j,k}(t) \quad (9)$$

2. Dans un second temps, on maximise, suivant \mathcal{X} , l'expression

$$\sum_{k=1}^M \sum_{t=1}^T \sum_{j=1}^{m_t} \log [p(z_j(t)|x_k(t))] w_{j,k}(t)$$

Le nouvel état mis à jour s'exprime sous la forme:

$$\forall t = 1 \dots T, k = 1 \dots M,$$

$$x_k(t) \in \text{Arg max}_{x_k} \sum_{j=1}^{m_t} \log [p(z_j(t)|x_k(t))] w_{j,k}(t) \quad (10)$$

La solution de cette dernière maximisation donne une solution unique dans le cas particulier d'observations linéaires. En règle générale, l'équation d'observation est non-linéaire. On présente dans le paragraphe suivant l'application de l'algorithme au sonar passif.

5.3 Exemple: Modèle non-linéaire des observations

Dans le domaine du sonar passif et en trajectographie par azimuts (TPA), on cherche à estimer la position et la vitesse de la source (cas d'une source évoluant à vitesse constante) dans un référentiel absolu à partir de relevés angulaires. Dans ce paragraphe, on restreint les observations à ne provenir que des sources. On ne prend donc pas en compte les fausses alarmes pour se concentrer sur le problème d'optimisation non-linéaire. L'équation d'observation de la mesure provenant du modèle k s'écrit alors:

$$z_j(t) = \beta_k(t) + \nu_k(t)$$

avec $\beta_k(t)$ une fonction non-linéaire de l'état $x_k(t)$

$$\beta_k = \arctan \left(\frac{x_k(t) - x_o(t)}{y_k(t) - y_o(t)} \right)$$

$x_o(t)$ et $y_o(t)$ désignent la position de l'observateur et $\nu_k(t)$ est un bruit blanc gaussien dont la variance, σ_k^2 est liée à l'estimation du gisement de la source k .

On suppose que les sources suivent un mouvement rectiligne uniforme, il s'agit alors d'estimer le vecteur d'état initial des sources que l'on notera X_k pour le modèle k . L'étape M de l'algorithme PMHT revient à trouver le vecteur d'état initial de chaque source séparément:

$$\forall k = 1 \dots M,$$

$$X_k^{i+1} \in \text{Arg max}_{X_k} \sum_{t=1}^T \sum_{j=1}^{m_t} \log [p(z_j(t)|X_k)] w_{j,k}(t)$$

avec $w_{j,k}(t) = \frac{\pi_k^i p(z_j(t)|x_k^i(t))}{p(z_j(t)|O^i(t))}$ et les probabilités d'assignation mises à jour

$$\pi_k^{i+1}(t) = \frac{1}{m_t} \sum_{j=1}^{m_t} w_{j,k}(t)$$

La densité de probabilité de la mesure j pour le modèle k s'exprime sous la forme:

$$p(z_j(t)|X_k) = \frac{1}{\sqrt{2\pi} \sigma_k} \exp \left[-\frac{1}{2\sigma_k^2} |z_j(t) - \beta_k(t)|^2 \right]$$

L'estimation du nouveau vecteur initial pour l'étape courante s'obtient en résolvant l'équation du gradient

$$\sum_{t=1}^T \frac{1}{\sigma_k^2} \frac{\partial \beta_k(t)}{\partial X_k} \sum_{j=1}^{m_t} (z_j(t) - \beta_k(t)) w_{j,k}(t) = 0$$

La solution de ce système d'équations non-linéaires est obtenu via un algorithme itératif du type Gauss-Newton.

6 Conclusion

On a utilisé l'approche originale au problème d'extraction introduite par R. Streit dans le cas d'un MAP récursif, pour estimer les paramètres des sources à l'aide du maximum de vraisemblance. Une telle approche tire parti de toute l'information contenue dans le batch en s'affranchissant des problèmes combinatoires liés à l'incertitude sur l'origine des mesures. On peut espérer obtenir des extracteurs plus performants. Les performances de ce type d'algorithme ont été étudiées sur les sorties simulées d'une chaîne sonar (sources multiples). Les résultats feront l'objet d'une publication future.

Références

- [1] T.E. Fortmann, Y. Bar-Shalom, M. Scheffe, "Sonar tracking of multiple targets using joint probabilistic data association", IEEE Journal of Oceanic Research, OE-8, pp 173-184, July 1983.
- [2] A. P. Dempster, N. M. Laird, D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm", J. Royal Statistical Society Ser. B, Vol. 39, pp. 1-38, 1977.
- [3] C. L. Morefield, "Application of 0-1 integer programming to multitarget tracking problems", IEEE Transactions on Automatic Control, Vol. AC-22, No. 3, pp. 302-311, June 1977.
- [4] D. B. Reid, "An Algorithm for Tracking Multiple Targets", IEEE Transactions on Automatic Control, Vol. AC-24, No. 6, pp. 843-854, December 1979.
- [5] R. A. Redner, H. F. Walker, "Mixture Densities, Maximum Likelihood and the EM Algorithm", Society for Industrial and Applied Mathematics, Vol. 26, No. 2, April 1984.
- [6] R. L. Streit, T. E. Luginbuhl, "Maximum Likelihood Method for Probabilistic Multi-Hypothesis Tracking", SPIE Int. Symposium, Signal and Data Processing of Small Targets 1994, SPIE Proceedings Vol. 2335-24, Orlando, FL, 5-7 April 1994.
- [7] C. F. Wu, "On the convergence of the EM algorithm", Annals of Statistics, Vol. 11, No. 1, pp. 95-103, 1983.