Estimation du centre et du rayon d'une hypersphère à l'aide d'une loi a priori de Von Mises-Fisher et d'un algorithme EM

Julien LESOUPLE[1], Barbara PILASTRE[2], Yoann ALTMANN[3], Jean-Yves TOURNERET[4,2]

[1]Laboratoire TELECOM, ENAC
7 Avenue Edouard Belin, 31500 Toulouse, France

[2]Laboratoire TéSA
7 Boulevard de la Gare, 31500 Toulouse, France

[3]Université Heriot-Watt
Campus d'Édimbourg, Édimbourg, EH14 4AS, Écosse

[4]Université de Toulouse, IRIT, INP-ENSEEIHT
2 Rue Charles Camichel, 31500 Toulouse, France

julien.lesouple@enac.fr, barbara.pilastre@tesa.prd.fr, y.altmann@hw.ac.uk,
jean-yves.tourneret@toulouse-inp.fr

Résumé – Cet article présente une extension d'un algorithme EM (*expectation maximization*) publié récemment par les auteurs permettant d'estimer conjointement le centre et le rayon d'une hypersphère avec les hyperparamètres d'un modèle statistique prenant en compte le fait que les observations sont localisées sur une partie de l'hypersphère. La méthode proposée repose sur l'ajout de variables latentes ayant une loi a priori de von Mises-Fisher. Ce modèle statistique permet d'exprimer la vraisemblance complète des données, dont l'espérance conditionnée aux données observées possède une distribution connue conduisant à un algorithme EM simple et efficace. Les performances de cet algorithme d'estimation sont évaluées à l'aide de de simulations effectuées dans un cas bi-dimensionnel avec des résultats prometteurs.

Abstract – This article introduces an extension of an EM algorithm (Expectation Maximization) published recently by the authors allowing to estimate jointly the center and the radius of an hypersphere as well as the statistical model hyperparameters acounting that the observations are located on a part of the hypersphere. The proposed method relies on the introduction of latent variables having a von Mises Fisher prior. This statistical model allows to express the complete data likelihood, which expectancy conditionned to the observed data has a known distribution resulting in a simple and efficient EM algorithm. The performances of this estimation algorithm are assessed through simulations performed in a bidimensinal case with promising results.

1 Introduction

L'estimation des paramètres d'une hypersphère (généralisation de la sphère en dimension quelconque) est d'un grand intérêt dans plusieurs applications comme la poursuite d'objets (tracking) [1–3], la robotique [4–6] ou le traitement d'images et la reconnaissance des formes [7–9]. Ce problème a récemment été formulé dans [10] avec l'introduction de variables latentes, définies comme des transformations affines de vecteurs aléatoires suivant une loi de von Mises-Fisher. La loi de von Mises-Fisher est une loi de probabilité définie sur l' hypersphère unité, qui est paramétrée par une direction moyenne et une concentration. L'estimation du rayon et du centre de l'hypersphère peut alors s'effectuer par un algorithme EM (*expectation maximization* décrit dans [10] dans le cas où les hyperparamètres de la loi a priori peuvent être spécifiés par l'utilisateur.

Cet article montre comment modifier l'algorithme de [10] pour estimer les hyperparamètres du modèle statistique utilisé pour l'estimation des paramètres d'une hypersphère. De plus, les équations de mise à jour de l'algorithme EM initial sont modifiées pour être plus explicites. Cet article est organisé comme suit. La section 2 rappelle comment le problème d'estimation des paramètres d'une d'hypersphère peut être formulé, ainsi que le principe de l'algorithme EM et les équations de mises à jour qui en découlent. La Section 3 évalue les performances de la méthode à travers différentes expérimentations. Quelques conclusions et perspectives sont indiquées dans la section 4.

2 Un Algorithme EM pour l'estimation des paramètres d'une hypersphère

2.1 Formulation du Problème

Considérons n observations bruitées $\boldsymbol{y}_i \in \mathbb{R}^d, i = 1, ..., n$ située autour d'une hypersphère de rayon $r > 0$ et de centre

$\boldsymbol{c} \in \mathbb{R}^d$. On suppose que les réalisations du bruit sont indépendantes et identiquement distribuées suivant une loi normale multivariée centrée et isotrope. On introduit alors des vecteurs latents $\boldsymbol{x}_i \in \mathcal{S}^{d-1}, i = 1, ..., n$, où $\mathcal{S}^{d-1}$ désigne l'hypersphère unité de $\mathbb{R}^d$ [10]. Ces vecteurs latents sont les vecteurs unitaires inconnus partant du centre et pointant vers le centre de la gaussienne générant l'observation, ce qui conduit au modèle suivant

$$\boldsymbol{y}_i = \boldsymbol{c} + r\boldsymbol{x}_i + \boldsymbol{e}_i, \tag{1}$$

où $\boldsymbol{e}_i \sim \mathcal{N}(\boldsymbol{0}_d, \sigma^2 \boldsymbol{I}_d)$ est l'erreur de modèle pour la i-ème observation, $\boldsymbol{0}_d$ est le vecteur nul de $\mathbb{R}^d$, $\sigma^2 > 0$ est la variance inconnue du bruit et $\boldsymbol{I}_d$ est la matrice identité de $\mathbb{R}^{d\times d}$.

Nous avons proposé dans [10] de munir les vecteurs $\boldsymbol{x}_i$ de lois a priori de von Mises-Fisher, i.e., $\boldsymbol{x}_i \sim \text{vMF}_d(\boldsymbol{x}_i; \boldsymbol{\mu}, \kappa)$ dont la densité est définie par

$$f_d(\boldsymbol{x}_i; \boldsymbol{\mu}, \kappa) = C_d(\kappa) \exp\left(\kappa \boldsymbol{\mu}^T \boldsymbol{x}_i\right) 1_{\mathcal{S}^{d-1}}(\boldsymbol{x}_i), \tag{2}$$

où $\boldsymbol{\mu} \in \mathbb{R}^d$ est la direction moyenne ($\|\boldsymbol{\mu}\|_2 = 1$), $\kappa \geq 0$ est la concentration, $1_{\mathcal{S}^{d-1}}(.)$ est la fonction indicatrice de $\mathcal{S}^{d-1}$, et $C_d(\kappa)$ est une constante de normalisation (rappelée dans [10]). Notons que la loi (2) se réduit à la loi uniforme sur l'hypersphère lorsque $\kappa = 0$. Cette loi est adaptée pour des applications LiDAR, dont la calibration peut se faire à l'aide de sphères [11]. En effet, dans ce cas, le faisceau LiDAR n'atteint qu'une partie de la sphère, ce qui donne un ensemble de points situées autour d'une partie de la sphère.

2.2 Vraisemblance et vraisemblance complète

D'après (1), la loi conditionnelle de $\boldsymbol{y}_i$ sachant $\boldsymbol{x}_i$ est

$$p(\boldsymbol{y}_i|\boldsymbol{x}_i, \boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{d/2}} \exp\left\{-\frac{\|\boldsymbol{y}_i - \boldsymbol{c} - r\boldsymbol{x}_i\|_2^2}{2\sigma^2}\right\}, \tag{3}$$

où $\boldsymbol{\theta}$ contient les paramètres du modèle à estimer, i.e., $\boldsymbol{\theta} = (r, \boldsymbol{c}^T, \sigma^2)^T$. Nous proposons dans cet article d'inclure également les hyperparamètres du problème κ et $\boldsymbol{\mu}$ dans ce vecteur pour obtenir finalement $\boldsymbol{\theta} = (r, \boldsymbol{c}^T, \sigma^2, \kappa, \boldsymbol{\mu}^T)^T$. En supposant que les erreurs de modèle $\boldsymbol{e}_i$ sont indépendantes, la vraisemblance marginale de ce modèle s'écrit

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{Y}) = \prod_{i=1}^{n} p(\boldsymbol{y}_i|\boldsymbol{\theta}) = \prod_{i=1}^{n} \int_{\mathcal{S}^{d-1}} p(\boldsymbol{y}_i, \boldsymbol{x}_i|\boldsymbol{\theta}) \mathrm{d}\boldsymbol{x}_i. \tag{4}$$

Malheureusement, l'estimateur du maximum de vraisemblance de $\boldsymbol{\theta}$ n'a pas d'expression explicite simple. Nous proposons donc de recourir à un algorithme EM [12] pour estimer le vecteur $\boldsymbol{\theta}$ comme dans [10]. On définit la vraisemblance complète

$$\mathcal{L}_c(\boldsymbol{\theta}; \boldsymbol{Y}, \boldsymbol{X}) = \prod_{i=1}^{n} p(\boldsymbol{y}_i, \boldsymbol{x}_i|\boldsymbol{\theta}). \tag{5}$$

De plus, en utilisant (3) et (2), on obtient

$$\begin{aligned} p(\boldsymbol{y}_i, \boldsymbol{x}_i|\boldsymbol{\theta}) &= p(\boldsymbol{y}_i|\boldsymbol{x}_i, \boldsymbol{\theta}) p(\boldsymbol{x}_i|\boldsymbol{\theta}), \\ &= \frac{C_d(\kappa)}{(2\pi\sigma^2)^{d/2}} \exp\left(-\frac{\|\boldsymbol{y}_i - \boldsymbol{c} - r\boldsymbol{x}_i\|_2^2}{2\sigma^2} + \kappa \boldsymbol{\mu}^T \boldsymbol{x}_i\right). \end{aligned} \tag{6}$$

Notez que la fonction indicatrice de $\boldsymbol{x}_i$ définie sur l'hypersphère unité a été omise dans (6) pour raccourcir cette équation, et il en sera ainsi tout au long de l'article même si on suppose bien entendu qu'on a $\boldsymbol{x}_i \in \mathcal{S}^{d-1}$.

2.3 Algorithme EM proposé

L'algorithme EM alterne entre deux étapes nommées *expectation* (E) et *maximization* (M), que l'on rappelle ci dessous pour l'itération $t + 1$:

1- Étape (E) : calcul de $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)})$. La fonction Q de l'algorithme EM est l'espérance du logarithme de la vraisemblance complète, par rapport à la loi des variables latentes sachant les observations et la valeur courante de l'estimation du vecteur des paramètres $\boldsymbol{\theta}^{(t)}$, i.e.,

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}) = \mathbb{E}_{\boldsymbol{X}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)}} \left[\log \mathcal{L}_c(\boldsymbol{\theta}; \boldsymbol{Y}, \boldsymbol{X})\right]. \tag{7}$$

2- Étape (M) : mise à jour des paramètres du modèle via

$$\boldsymbol{\theta}^{(t+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)}). \tag{8}$$

La log-vraisemblance complète se calcule avec (5) et (6)

$$\begin{aligned} \log \mathcal{L}_c(\boldsymbol{\theta}; \boldsymbol{Y}, \boldsymbol{X}) &= K - \frac{nd}{2} \log \sigma^2 + n \log C_d(\kappa) \\ &\quad - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \left(\|\boldsymbol{y}_i - \boldsymbol{c}\|_2^2 + r^2\right) + \sum_{i=1}^{n} \kappa_i \boldsymbol{\mu}_i^T \boldsymbol{x}_i, \end{aligned} \tag{9}$$

où K est un terme indépendant de $\boldsymbol{\theta}$, et

$$\kappa_i = \frac{\|r(\boldsymbol{y}_i - \boldsymbol{c}) + \sigma^2 \kappa \boldsymbol{\mu}\|_2}{\sigma^2}, \tag{10}$$

$$\boldsymbol{\mu}_i = \frac{r(\boldsymbol{y}_i - \boldsymbol{c}) + \sigma^2 \kappa \boldsymbol{\mu}}{\|r(\boldsymbol{y}_i - \boldsymbol{c}) + \sigma^2 \kappa \boldsymbol{\mu}\|_2}. \tag{11}$$

La loi de $\boldsymbol{X}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)}$ peut se calculer comme suit

$$p(\boldsymbol{X}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)}) = \prod_{i=1}^{n} p(\boldsymbol{x}_i|\boldsymbol{y}_i, \boldsymbol{\theta}^{(t)}). \tag{12}$$

En utilisant (6), (10), (11) et (2), on obtient

$$\begin{aligned} p(\boldsymbol{x}_i|\boldsymbol{y}_i, \boldsymbol{\theta}) &\propto p(\boldsymbol{y}_i, \boldsymbol{x}_i|\boldsymbol{\theta}), & (13) \\ &\propto f_d(\boldsymbol{x}_i; \boldsymbol{\mu}_i, \kappa_i), & (14) \end{aligned}$$

où $\propto$ signifie "proportionnel à". On reconnaît alors une loi de von Mises-Fisher de paramètre κ_i et $\boldsymbol{\mu_i}$, dont la moyenne vaut

$$A_d(\kappa_i) = \frac{I_{d/2}(\kappa_i)}{I_{d/2-1}(\kappa_i)}, \tag{15}$$

où $I_\nu(.)$ désigne la fonction de Bessel modifiée du premier type de paramètre ν [13, Chap. 10.25]. On en déduit alors

$$\mathbb{E}_{\boldsymbol{X}|\boldsymbol{Y}, \boldsymbol{\theta}^{(t)}}[\boldsymbol{x}_i] = A_d\left(\kappa_i^{(t)}\right) \boldsymbol{\mu}_i^{(t)}, \tag{16}$$

où $\kappa_i^{(t)}$, et $\boldsymbol{\mu}_i^{(t)}$ peuvent être calculés à l'itération t en utilisant (10) et (11) pour la valeur courante $\boldsymbol{\theta}^{(t)}$. On peut alors remplacer les expressions obtenues dans (9), pour maximiser la fonction $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t)})$ par rapport à $\boldsymbol{\theta}$, ce qui conduit à

$$r^{(t+1)} = \frac{\overline{\boldsymbol{\alpha}^T \boldsymbol{y}}_t - \overline{\boldsymbol{\alpha}}_t^T \overline{\boldsymbol{y}}_t}{1 - \overline{\boldsymbol{\alpha}}_t^T \overline{\boldsymbol{\alpha}}_t}, \tag{17}$$

$$\boldsymbol{c}^{(t+1)} = \overline{\boldsymbol{y}} - r^{(t+1)} \overline{\boldsymbol{\alpha}}_t, \tag{18}$$

$$\begin{aligned} d\sigma^{2(t+1)} &= \overline{\|\boldsymbol{y}\|_2^2} + \|c^{(t+1)}\|_2^2 + {r^{(t+1)}}^2 \\ &\quad - 2\left\{\boldsymbol{c}^{(t+1)T} \overline{\boldsymbol{y}}_t + r^{(t+1)} \left[\overline{\boldsymbol{\alpha}^T \boldsymbol{y}}_t - \overline{\boldsymbol{\alpha}}_t^T \boldsymbol{c}^{(t+1)}\right]\right\} \end{aligned} \tag{19}$$

$$\kappa^{(t+1)} = A_d^{-1}(\|\overline{\boldsymbol{u}}_t\|_2) \tag{20}$$

$$\boldsymbol{\mu}^{(t+1)} = \frac{\overline{\boldsymbol{u}}_t}{\|\overline{\boldsymbol{u}}_t\|_2} \tag{21}$$

avec

$$\overline{\boldsymbol{y}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{y}_i \tag{22}$$

$$\overline{\boldsymbol{\alpha}}_t = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\alpha}_i^{(t)} \tag{23}$$

$$\overline{\|\boldsymbol{y}\|_2^2} = \frac{1}{n}\sum_{i=1}^{n} \|\boldsymbol{y}_i\|_2^2 \tag{24}$$

$$\boldsymbol{\alpha}_i^{(t)} = A_d(\kappa_i^{(t)})\boldsymbol{\mu}_i^{(t)} \tag{25}$$

$$\overline{\boldsymbol{\alpha}^T \boldsymbol{y}}_t = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{y}_i^T \boldsymbol{\alpha}_i^{(t)} \tag{26}$$

Remarquons que les quantités surmontées d'une barre correspondent à des moyennes empiriques, et que $\alpha_i^{(t)}$ correspond à la moyenne d'une loi de von Mises-Fisher de paramètres $\kappa_i^{(t)}$ et $\boldsymbol{\mu}_i^{(t)}$. Les équations de mise à jour (20) et (21) ont été obtenues en utilisant les expressions des estimateurs du maximum de vraisemblance pour les paramètres d'une distribution de von Mises-Fisher [14, Chap. 10.3.1]. La fonction $A_d^{-1}(\cdot)$ n'a pas d'expression littérale, mais peut-être calculée avec un algorithme itératif [15]. Comparé à l'algorithme original [10], nous avons rajouté les équations de mises à jour de κ (20) et de $\boldsymbol{\mu}$ (21) (alors que l'algorithme original suppose ces paramètres connus) et les équations (17), (18) pour r et $\boldsymbol{c}$ sont découplées et ne font pas intervenir d'invertion matricielle.

3 Performances de l'estimateur EM

Dans toutes les expérimentations, l'algorithme EM est initialisé en utilisant l'algorithme IML [16] afin d'obtenir une première valeur pour le rayon et le centre de l'hypersphère considérée. Ensuite, les autres paramètres (variance, et hyperparamètres de la loi de von Mises-Fisher) sont estimés par maximum de vraisemblance [17, Chap. 7], i.e., $\hat{\kappa} = A_d^{-1}(\|\bar{\boldsymbol{y}}\|_2)$, $\hat{\sigma}^2 = \frac{1}{d}\left(\|\bar{\boldsymbol{y}} - \hat{\boldsymbol{c}}\|_2^2 - \hat{r}^2\right)$ et $\hat{\mu} = \frac{\bar{\boldsymbol{y}}}{\|\bar{\boldsymbol{y}}\|_2}$, où $\bar{\boldsymbol{y}}$ est la moyenne empirique des n observations, et $A_d^{-1}(\cdot)$ a été définie précédemment.

3.1 Exemple 2D

Dans cette première expérience, nous considérons un cercle ($d = 2$) de paramètres $r = 6$, $\boldsymbol{c} = (1,1)^T$, $\sigma^2 = 1$, $\kappa \in \{0, 5\}$ et $\boldsymbol{\mu} = (1,1)^T$. Dans la Figure 1, nous avons représenté, pour chacune des deux valeurs de κ, les observations, ainsi que le cercle théorique, le cercle estimé par l'algorithme IML (point de départ pour l'algorithme EM) et le cercle obtenu avec l'algorithme proposé. On remarque que les deux algorithmes semblent converger vers une solution satisfaisante quand $\kappa = 0$ (distribution uniforme). En revanche quand $\kappa = 5$ (une direction de l'espace est privilégiée), la méthode proposée est clairement plus performante, car elle exploite le fait qu'on estime cette direction privilégiée.

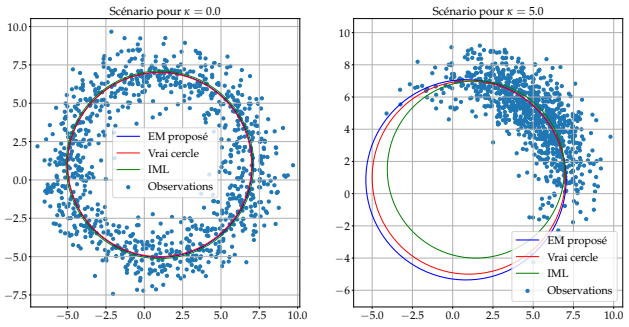


FIGURE 1 – Observations, vérité terrain et cercles estimés par la méthode proposée et la méthode IML pour deux valeurs de κ ($\kappa = 0$ à gauche et $\kappa = 5$ à droite).

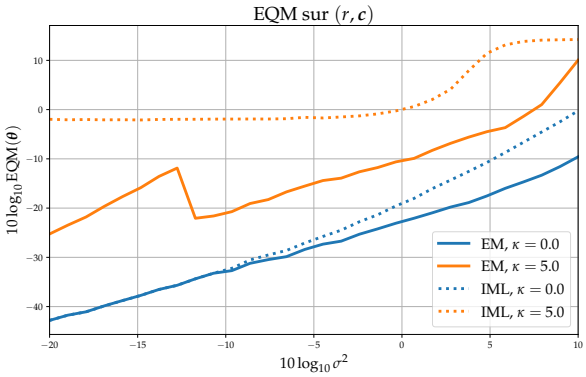


FIGURE 2 – EQM de $(r, \boldsymbol{c}^T)^T$ (paramètres d'intérêt du cercle) pour différentes valeurs de κ.

3.2 Simulations de Monte-Carlo

Pour avoir une idée de la performance globale de l'algorithme présenté, nous avons effectué des simulations de Monte-Carlo, en répétant l'expérience précédente $N_{\text{MC}} = 500$ fois afin de pouvoir estimer l'erreur quadratique moyenne (EQM) des différents paramètres du problème. Nous avons représenté dans la Figure 2 l'EQM des paramètres d'intérêt r et $\boldsymbol{c}$ du cercle pour les deux valeurs de κ. On remarque que dans les deux cas, l'algorithme présenté est plus performant que l'algorithme de référence IML. Cela avait déjà été montré dans [10], dans le cas d'hyperparamètres $(\kappa, \boldsymbol{\mu})$ connus. On observe ici que ce gain de performance est conservé lorsque ces hyperparamètres sont estimés. Dans la figure 3, nous avons représenté l'EQM de l'hyperparamètre de concentration. Nous remarquons que dans le cas $\kappa = 0$, les deux estimateurs se comportent de manière similaire. Même si l'estimation via IML semble légèrement meilleure pour des forts niveaux de bruit, les valeurs d'EQM obtenues avec l'algorithme EM restent satisfaisantes (≈ -20 dB), avec l'avantage de fournir de meilleures estimations des paramètres d'intérêt du cercle. Finalement, nous avons représenté dans la Figure 4 l'EQM du vecteur direction moyenne (uniquement pour $\kappa > 0$, car pour $\kappa = 0$, la loi de von Mises-Fisher est une loi uniforme et il n'y a donc pas de direction privilégiée). On remarque une fois encore que l'algorithme proposé offre une meilleure estimation.

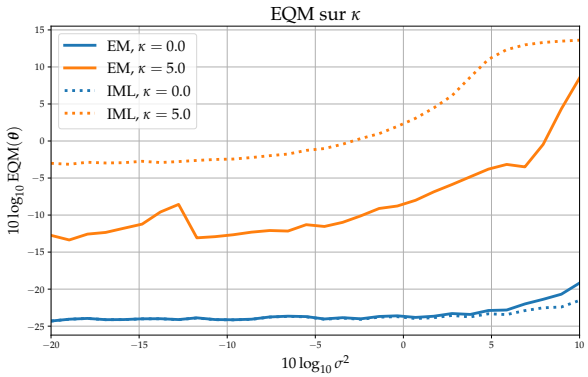


FIGURE 3 – EQM de l'hyperparamètre κ.

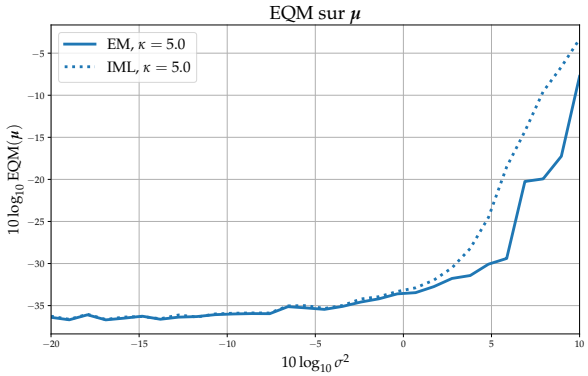


FIGURE 4 – EQM de l'hyperparamètre μ.

4 Conclusion

Nous avons présenté dans cet article un algorithme permettant l'estimation des paramètres d'une hypersphère, en utilisant des observations proches d'une direction privilégiée inconnue par l'utilisateur. L'algorithme proposé repose sur l'introduction de variables latentes ayant une distribution de von Mises-Fisher dont les paramètres (appelés hyperparamètres) sont estimés conjointement aux paramètres d'intérêt de l'hypersphère. Cette estimation des hyperparamètres n'a pas un impact significatif sur les performances de l'algorithme EM, ce qui est très intéressant. Nous travaillons actuellement sur l'étude des performances optimales des estimateurs des paramètres d'hypersphères avec le modèle statistique proposé dans cet article.

Références

[1] M. Baum and U. D. Hanebeck, "Random Hypersurface Models for Extended Object Tracking," in *Proc. IEEE Int. Symp. Sign. Process. Inf. Tech. (ISSPIT 2009)*, Ajman, United Arab Emirates, Dec. 2009, pp. 178–183.

[2] Niklas Wahlström and Emre Özkan, "Extended Target Tracking Using Gaussian Processes," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 1, pp. 4165–4178, May 2015.

[3] D. Epstein and D. Feldman, "Sphere Fitting with Applications to Machine Tracking," *Algorithms*, vol. 13, no. 8, pp. 177, July 2020.

[4] F. Sandoval, "An Algorithm for Fitting 2-D Data on the Circle : Applications to Mobile Robotics," *IEEE Signal Process. Lett.*, vol. 15, pp. 127–130, Jan. 2008.

[5] F. Bonin-Font, A. Ortiz, and G. Oliver, "Visual Navigation for Mobile Robots : A Survey," *J. Intell. Robot. Syst.*, vol. 53, pp. 263–296, Nov. 2008.

[6] D. Epstein and D. Feldman, "Quadcopter Tracks Quadcopter via Real Time Shape Fitting," *IEEE Robot. Autom. Lett.*, vol. 3, pp. 544–550, Jan. 2018.

[7] D. Lin and C. Yang, "Real-Time Eye Detection Using Face-Circle Fitting and Dark-Pixel Filtering," in *Proc. IEEE Int. Conf. on Multimedia ans Expo (ICME'2004)*, Taipei, Taiwan, June 2004, pp. 1167–1170.

[8] L. Pan, W. Chu, J. M. Saragih, F. De la Torre, and M. Xie, "Fast and Robust Circular Object Detection With Probabilistic Pairwise Voting," *IEEE Sign. Process. Let.*, vol. 18, no. 11, pp. 639–642, Sept. 2011.

[9] A. Geiger, P. Lenz, and R. Urtasun, "Are We Ready for Autonomous Driving ? The Kitti Vision Benchmark Suite," in *Proc. IEEE Int. Conf. Conf. Comput. Vis. Pattern Recognit. (CVPR'2012)*, Providence, RI, USA, Jun. 2012, pp. 3354–3361.

[10] Julien Lesouple, Barbara Pilastre, Yoann Altmann, and Jean-Yves Tourneret, "Hypersphere Fitting from Noisy Data Using an EM Algorithm," *IEEE Signal Process. Lett.*, vol. 28, pp. 314–318, 1 2021.

[11] Minghao Ruan and Daniel Huber, "Calibration of 3D Sensors Using a Spherical Target," in *Proc. 2nd International Conference on 3D Vision*, Tokyo, Japan, Dec. 2014, pp. 187–193.

[12] A.P. Dempster, N.M. Laird, and Do. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *J. R. Stat. Soc. Series B*, vol. 39, no. 1, pp. 1–38, 1977.

[13] Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert, and Charles W. Clark, *NIST Handbook of Mathematical Functions*, Cambridge University Press, 2010.

[14] Kanti V. Mardia and Peter E. Jupp, *Directional Statistics*, John Wiley & Sons, Inc., 1999.

[15] Suvrit Sra, "A Short Note on Parameter Approximation for von Mises-Fisher Distributions : and a Fast Implementation of $I_s(x)$," *Computational Statistics*, vol. 27, no. 1, pp. 177–190, March 2012.

[16] W. Li, J. Zhong, T. A. Gulliver, B. Rong, R. Hu, and Y. Qian, "Fitting Noisy Data to a Circle : A Simple Iterative Maximum Likelihood Approach," in *Proc. IEEE Int. Conf. on Communications*, Kyoto, Japan, July 2011, pp. 1 – 5.

[17] Steven M. Kay, *Fundamentals of Statistical Signal Processing : Estimation Theory*, Prentice-Hall, Inc., USA, 1993.