



HAL
open science

Vers la navigation référencée terrain par mise en correspondance de cartes bathymétriques

Laurence Lucido

► **To cite this version:**

Laurence Lucido. Vers la navigation référencée terrain par mise en correspondance de cartes bathymétriques. RR-3211, INRIA. 1997. inria-00073478

HAL Id: inria-00073478

<https://inria.hal.science/inria-00073478>

Submitted on 24 May 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Unité de recherche INRIA Lorraine, technopôle de Nancy-Brabois, 615 rue du jardin botanique, BP 101, 54600 VILLERS-LÈS-NANCY
Unité de recherche INRIA Rennes, IRISA, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unité de recherche INRIA Rocquencourt, domaine de Voluceau, Rocquencourt, BP 105, LE CHESNAY Cedex
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN 93,
06902 SOPHIA-ANTIPOLIS Cedex

Éditeur

INRIA, Domaine de Voluceau, Rocquencourt, BP 105 LE CHESNAY Cedex (France)

ISSN 0249-6399

- [48] O. Faugeras, F. Lustman et G. Toscani. *Calcul du mouvement et de la structure à partir de Points et de Droites*. Rapport de Recherches 661, INRIA Sophia Antipolis, Mai 1987.
- [49] X. Pennec et J.P. Thirion. *Validation of 3D registration methods based on points and frames*. Rapport de Recherches, INRIA Sophia Antipolis, Janvier 1995.
- [50] J.-M. Chassery et A. Montanvert. *Géométrie Discrète en Analyse d'Images*. Hermès, 1991.

-
- [37] O. Bergem. *Estimation of the measurement covariance matrix in a Kalman filter*. Proc. of the 4th Int. Workshop on Time-Varying Image Processing and Moving Object Recognition, Florence, Italie, Juin 1993.
- [38] R. Cristi. *Sensor based navigation of an autonomous underwater vehicle*.
- [39] J.P. Allenou. *Essais des Systèmes multi-faisceaux EM1000-EM12*, Navire Jean Charcot. Genavir, mai 1992.
- [40] J.M. Augustin, C. Edy, S. Savoye et E. LeDrezen. *Sonar Mosaic Computation from Multibeam Echo Sounder*. Oceans94 pp. 433-438, 13-16 Sept. 1994, Brest France.
- [41] I. J. Cox. *Blanche : Position Estimation for an Autonomous Robot Vehicule*. Autonomous Robot Vehicules, Springer-Verlag.
- [42] Z. Zhang et O. Faugeras. *3D Dynamic Scene Analysis : A Stereo Based Approach*. Springer, Berlin, Heidelberg, 1992.
- [43] Z. Zhang. *Le problème de la mise en correspondance : l'état de l'art*. Rapport de Recherches 2146, INRIA Sophia Antipolis, Décembre 1993.
- [44] L. G. Brown. *A survey of Image Registration Techniques*. ACM Computing Survey. Vol.24 No4, pp. 325-376.
- [45] Z. Zhang et O. Faugeras. *Estimation of Displacements from two 3D frames obtained from stereo*. Rapport de Recherches 1440, INRIA Sophia Antipolis, Juin 1991.
- [46] Z. Zhang et O. Faugeras. *Three-dimensional motion computation and object segmentation in a long sequence of stereo frames*. Rapport de Recherches 1438, INRIA Sophia Antipolis, Juillet 1991.
- [47] S. Lee et Y. Kay. *A Kalman filter approach for accurate 3D motion estimation from a sequence of stereo images*. CVGIP : Image Understanding, Vol. 54 No 2, pp. 244-258, Septembre 1991.

-
- [25] A. Chéhikian. *Algorithmes optimaux pour la génération de pyramides d'images passe-bas et laplaciennes*. Traitement du Signal Vol 9, No4, 1992.
- [26] P.J. Burt. *The pyramid as a structure for efficient computation*. Multi-resolution Image Processing and Analysis, Editor A. Rosenfeld, Berlin, 1984.
- [27] A. Witkin. *Scale-Space Filtering*. Proc. 8th Int. Joint. Conf. Art. Intell., pp. 1019-1022, Aout 1983.
- [28] T. Lindeberg et B. M. ter Haar Romeny. *Linear Scale-Space*. Geometry-Driven Diffusion in Computer Vision, 1994.
- [29] T. Lindeberg. *Discrete Derivative Approximations with Scale-Space Properties : A Basis for Low-Level Feature Extraction*. Journal of Mathematical Imaging and Vision. Vol.3, No4, pp. 349-376, 1993.
- [30] F. Glazer. *Scene matching by hierarchical correlation*. 1983.
- [31] B. Sabata et J.K. Aggarwal. *Estimation of Motion from a pair of range images : a review*. CVGIP : Image Understanding, Vol. 54 No 3, pp. 309-324, Novembre 1991.
- [32] D.S. Kalivas et A.A. Sawchuk. *A region matching motion estimation algorithm*. CVGIP : Image Understanding, Vol. 54, No. 2, pp. 275-288, Septembre 1991.
- [33] J.-M. Odobez et P. Bouthémy. *Robust multiresolution estimation of parametric motion models in complex image sequences*. 7th European Conf. on Signal Processing, Edinburgh, Septembre 1994.
- [34] J.-M. Odobez et P. Bouthémy. *Robust multiresolution estimation of parametric motion models applied to complex scenes*. IRISA Publication Interne No 788, Janvier 1994.
- [35] R.L. Marks, M.J. Lee et S.M. Rock. *Using Visual Sensing for Control of an Underwater Robotic Vehicule*.
- [36] O. Bergem. *A multibeam sonar based positioning system for AUV*.

Bibliographie

- [15] R.A. Samy et C. Bozzo. *Extension de la Transformation de Hough au traitement d'une séquence d'images*. Premier Colloque Image, Biarritz, Mai 1984.
- [16] G. Giraudon. *Redésignation sur séquence d'images*. Premier Colloque Image, Biarritz, Mai 1984.
- [17] J.J. Rodriguez et J.K. Aggarwal. *Matching Aerial Images to 3D Terrain Maps*. IEEE Trans. on PAMI, vol. 12, No 12, Décembre 1990.
- [18] J.J. Rodriguez et J.K. Aggarwal. *Navigation Using Image Sequence Analysis and 3D Terrain Matching*. Workshop on Interpretation of 3D Scenes, Novembre 1989, Austin (Texas).
- [19] J.J. Rodriguez. *Terrain Matching using image sequence analysis*. Ph.D. 1990, Univ. of Texas, Austin.
- [20] B. Kamgar-Parsi, J.L. Jones et A. Rosenfeld. *Registration of multiple overlapping range images : Scenes without distinctive features*. IEEE PAMI Vol 13, No 9, septembre 1991.
- [21] D.B. Goldgof, T.S. Huang et H. Lee. *Feature Extraction and Terrain Matching*. CVPR/IEEE, conf. Ann Arbor (USA), Juin 1988.
- [22] Z. Zhang. *Recalage de deux nuages de points 3D*. Traitement du Signal Vol 10, No4, 1993.
- [23] Z. Zhang. *Itérative point matching for registration of free-form curves and surfaces*. International Journal of Computer Vision, Vol 13, No 2, pp. 119-152, 1994.
- [24] J.-M. Jolion. *Analyse d'images : le modèle pyramidal*. Traitement du Signal Vol 7, No1, 1990.

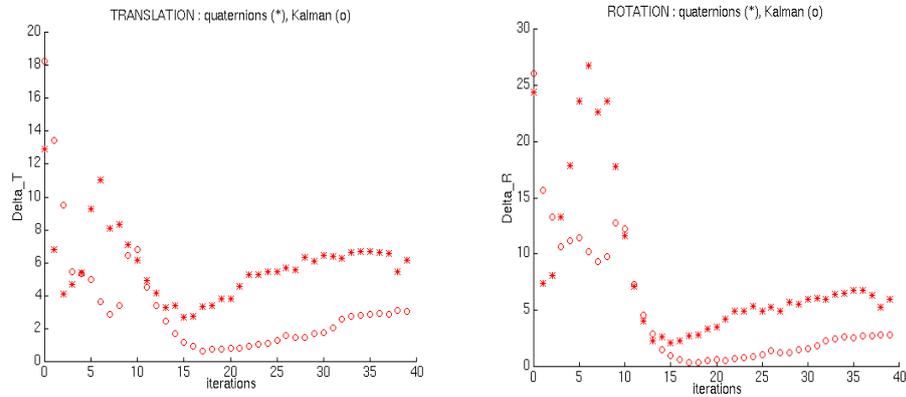


Figure 14. *Comparaison de la convergence des 2 approches (test 3)*

Les résultats sont sensiblement équivalents, malgré un temps de calcul plus important pour le filtrage de Kalman. Toutefois, le filtre de Kalman est intéressant car il est itératif (il peut être appliqué à de nouvelles données), il tient compte de l'incertitude sur les données (il est robuste au bruit sur les données) ; de plus, il fournit en sortie la matrice de covariance de l'état estimé (en d'autres termes, il fournit une information sur la confiance dans le résultat).

Remerciements. L'auteur tient à remercier ici son laboratoire d'accueil, le laboratoire de Robotique et d'Intelligence Artificielle de l'IFREMER - centre de Toulon, et son responsable, Vincent Rigaud. Il adresse ses remerciements à Rachid Deriche et Zhengyou Zhang, projet RobotVis de l'INRIA, pour leur implication au projet de recherche et leur précieuse collaboration au cours de cette étude.

test	Estimée du mouvement par quaternions	Différence finale	Nombre Itérations	Temps de calcul
1	[6.639549e-03 -8.151732e-03 2.096024e-02 2.179576e-02 -1.678528e-01 -3.936255e-02]	rotation : 1.34 degré translation : 17 cm	40	143
2	[-1.022265e-02 -6.817966e-03 6.414718e-02 5.362147e-02 -3.399910e-01 -2.713074e-02]	rotation : 3.74 degrés translation : 34.5 cm	40	144
3	[4.190397e-03 -2.562330e-03 2.018841e-02 1.450378e-02 -1.224675e-01 -9.507817e-03]	rotation : 1.18 degrés translation : 12.3 cm	80	390

Tableau 6. Résultats après minimisation par la méthode des quaternions

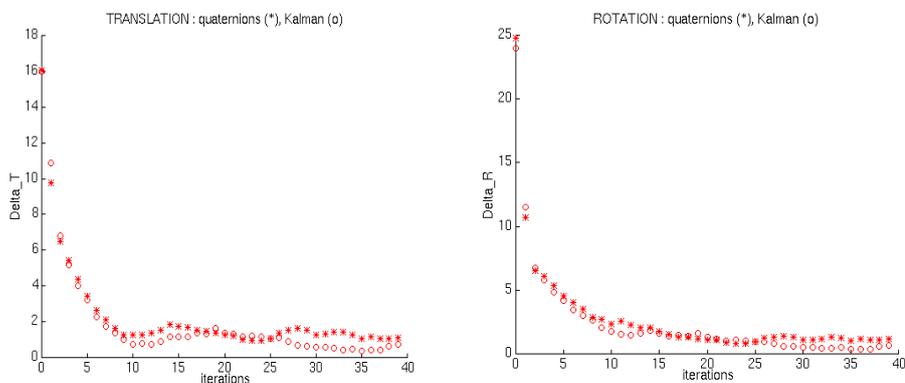


Figure 13. Comparaison de la convergence des 2 approches (test 2) avec $\Delta X = \|X_i - X_{i-1}\| / \|X_i\|$

test	r0	t0	Différence initiale
1	[0 0 0.17]	[0 -1.5 0]	rotation : 10 degrés translation : 1.5 m
2	[0 0 0.35]	[0.5 -2 0.2]	rotation : 20 degrés translation : 2.07 m
3	[0.35 0 0.17]	[-0.5 -2.5 0.2]	rotation : 22.3 degrés translation : 2.56 m

Tableau 4. Données initiales

test	Estimée du mouvement par Kalman	Différence finale	Nombre Itérations	Temps de calcul*
1	[6.329228e-03 -7.848704e-03 2.082458e-02 2.072145e-02 -1.665686e-01 -3.764577e-02]	rotation : 0.6 degré translation : 17 cm	40	232
2	[-9.763798e-03 -7.080189e-03 6.244803e-02 5.125999e-02 -3.320311e-01 -2.827827e-02]	rotation : 3.64 degrés translation : 33.5 cm	40	237
3	[4.378456e-03 -2.349055e-03 1.934871e-02 1.372833e-02 -1.187207e-01 -8.479600e-03]	rotation : 1.14 degrés translation : 11 cm	80	662

Tableau 5. Résultats après filtrage de Kalman (* temps CPU en secondes sur station de travail SUN Sparc).



Figure 12. Les scènes de rocher (avant et après le mouvement de la caméra)

test	r0	t0	δr0	δt0
1	[0.3 0.3 0.3]	[100 -100 100]	37 %	58.67 %
2	[0 0 0]	[100 -100 100]	100 %	58.67 %
3	[0.3 0.3 0.3]	[0 0 0]	37 %	100 %
4	[0 -0.1 0.8]	[-200 -200 -200]	87 %	164 %

Tableau 2. Données Initiales

test	δr	δt	r estimée	t estimée	itérations
1	0.04408%	0.5451 %	[0.399 0.199 0.499]	[200.43 -149.35 298.01]	2
2	divergence	0.5820 %	divergence	[200.42 -148.86 298.08]	6
3	0.04415 %	0.54276 %	[0.399 0.199 0.199]	[192.67 -149.27 260.96]	2
4	0.04389 %	0.5544 %	[0.399 0.199 0.499]	[200.39 -149.03 298.10]	3

Tableau 3. Résultats

On constate que la linéarisation de l'équation du système conduit à une très forte interaction entre les paramètres du mouvement ; notamment, on constate l'influence des paramètres de la rotation sur l'ensemble du résultat final. A ce sujet, [49] développe un formalisme original pour modéliser l'erreur sur les transformations rigides : celle-ci n'est plus additive, mais elle peut être vue comme la *composition* de la valeur réelle de la transformation avec une erreur : si p est le vecteur réel et p' le vecteur mesuré, alors $p' = p_o e$, où e est un petit mouvement rigide autour de l'identité. Toutefois, une telle modélisation est applicable dans le cas où les données mises en correspondance sont des *structures* définies par l'association de trièdres directs aux points 3D.

10.2 sur des données réelles

Entre les deux vues (figure 12) et d'après [22], le mouvement est de 30 degrés en rotation et 3.75 m en translation. Un premier recalage manuel a été effectué sur ces vues, de telle sorte que l'estimée finale attendue est zéro.

La matrice de covariance initiale du mouvement est :

$$S_0 = \text{diag}(1, 1, 1, 10^8, 10^8, 10^8)$$

10.1 sur des données synthétiques

Nous avons généré un fichier de données (50 points 3D) auxquelles nous avons appliqué une transformation tri-dimensionnelle (rotation puis translation). Les données ainsi obtenues sont bruitées par un bruit gaussien de moyenne nulle et de variance 10. Il n'y a pas d'appariement incorrect.

Les composantes du vecteur r sont exprimées en radians, celles de t en cm.

Le mouvement correct, qui doit être le résultat final du filtre de Kalman, est :

$$r = [0.4 \ 0.2 \ 0.5]^T \text{ et } t = [200 \ -150 \ 300]^T$$

ce qui correspond à une rotation de 38 degrés autour d'un axe de vecteur directeur r et passant par l'origine du repère, et une translation de 3.90 m.

La matrice de covariance initiale de l'état (choisie grande pour modéliser l'incertitude sur l'estimation initiale du mouvement) est :

$$S_0 = \text{diag}(1, 1, 1, 10^8, 10^8, 10^8)$$

Le filtre de Kalman étendu utilise une approximation au premier ordre de l'équation non-linéaire du système. Dans le cas où l'estimation initiale est très différente du mouvement réel, on voit le filtre converger vers des minima locaux, souvent erronés. Pour pallier à cet effet et le réduire dans certains cas, on itère plusieurs fois le filtre sur toutes les données : à chaque itération, on calcule une nouvelle estimée \hat{p}_i du mouvement et sa matrice de covariance S_i en choisissant comme estimée initiale l'estimée \hat{p}_{i-1} obtenue à l'estimation précédente mais en conservant la matrice de covariance S_0 .

On définit les erreurs (%) en rotation δr et en translation δt par :

$$\delta r = \frac{\|r - \hat{r}\|}{\|r\|} \times 100 \text{ et } \delta t = \frac{\|t - \hat{t}\|}{\|t\|} \times 100 \quad (\text{EQ 73})$$

Les tableaux 2 et 3 ci-dessous résument quelques tests effectués sur des données simulées, en choisissant différentes estimations initiales. On constate que l'algorithme diverge (cas No2) lorsque l'estimée initiale des paramètres de rotation est très différente du mouvement correct. Toutefois on obtient de bons résultats (cas No4) même lorsque le choix des paramètres initiaux de mouvement sont fort éloignés du cas réel.

$$\Lambda_{\chi'} = \begin{bmatrix} \Lambda_M & 0 \\ 0 & \Lambda_{M'} \end{bmatrix} \text{ pour tout couple } (M, M').$$

On choisira en pratique : $\Lambda_M = \Lambda_{M'} = \sigma^2 I_3$.

Par suite, la matrice de covariance du bruit Λ_ξ s'exprime, d'après l'équation 8, par :

$$\Lambda_\xi = \frac{\partial f(\chi', p)}{\partial \chi'} \begin{bmatrix} \Lambda_M & 0 \\ 0 & \Lambda_{M'} \end{bmatrix} \frac{\partial f(\chi', p)^T}{\partial \chi'}$$

$$\Lambda_\xi = R\Lambda_M R^T + \Lambda_{M'} \quad (\text{EQ 72})$$

9.3 Condition de convergence

La condition de terminaison d'itération est définie dans [23] comme la variation de l'estimée du mouvement entre deux itérations successives. La variation en translation après l'itération I est définie par :

$$\delta t = \frac{\|t_I - t_{I-1}\|}{\|t_I\|}$$

De même, la variation en rotation est définie par :

$$\delta r = \frac{\|r_I - r_{I-1}\|}{\|r_I\|}$$

On considère que le filtre a convergé lorsque δt et δr sont toutes les deux inférieures à 1%, ou bien lorsque le nombre d'itérations a atteint un seuil préfixé (50, par exemple).

10 Les résultats ...

Nous allons présenter dans ce paragraphe les résultats obtenus sur des données synthétiques pour valider l'algorithme d'estimation de mouvement par filtre de Kalman, puis sur des données réelles. C'est sur ce dernier point que nous pourrons comparer la procédure de minimisation par la méthode du quaternion et le filtrage de Kalman pour l'estimation du mouvement 3D.

- Equation de mesure : $f(\chi', p) = [M' - RM - t] = 0$

Un développement de Taylor au premier ordre autour du point (χ_i, \hat{p}_{i-1}) donne la nouvelle équation de mesures : $Y_i = A_i \cdot p + \xi_i$.

Par suite, on mesure Y car on connaît f et l'estimée précédente du paramètre p et on observe χ , la version bruitée de χ' . En outre, on a N mesures χ'_i vérifiant l'équation de mesure et l'estimée initiale \hat{p}_0 de p de matrice de covariance $\Lambda_{\hat{p}_0}$. On peut alors appliquer un filtre de Kalman étendu dont les équations sont données au paragraphe 2.2.2. Lorsque toutes les mesures sont utilisées, le paramètre p est connu par son estimée a posteriori \hat{p}_N et sa matrice de covariance $\Lambda_{\hat{p}_N}$.

9.1 Calcul des dérivées de $f(\chi', p)$

- par rapport au vecteur χ'

$$\frac{\partial f(\chi', p)}{\partial \chi'} = [-R \ I_3] \quad (\text{EQ 69})$$

- par rapport au vecteur p :

$$\frac{\partial f(\chi', p)}{\partial p} = - \left[\frac{\partial RM}{\partial r} \ I_3 \right] \quad (\text{EQ 70})$$

avec :

$$\begin{aligned} \frac{\partial RM}{\partial r} = & \frac{\cos \theta - f(\theta)}{\theta^2} (r \wedge M) r^T + \frac{\sin \theta - 2\theta g(\theta)}{\theta^3} (r \wedge (r \wedge M)) r^T - f(\theta) \tilde{M} \\ & + g(\theta) \cdot \{ -[(r \wedge \tilde{M}) + (rM)I_3 - Mr^T] \} \end{aligned} \quad (\text{EQ 71})$$

9.2 Le traitement des incertitudes

Le filtre de Kalman nécessite la connaissance de la matrice de covariance sur chacune des mesures :

$f(\theta) = \sin\theta/\theta$ et $g(\theta) = (1 - \cos\theta)/\theta^2$, sachant que $e^{\tilde{r}}$ est une notation pour l'exponentielle de la matrice \tilde{r} .

Par suite, nous représenterons désormais un mouvement rigide 3D (R, t) par un vecteur p réel de dimension 6 :

$$p = \begin{bmatrix} r \\ t \end{bmatrix} \quad (\text{EQ 68})$$

9 L'estimation du mouvement 3D

Le but de notre étude est de déterminer les paramètres de mouvement (rotation et translation) entre deux cartes de profondeur successives. On dispose, en entrée de l'algorithme, d'un ensemble de couples de points 3D appariés. On suppose que le mouvement entre les deux vues est petit, ie. on a une estimation initiale du mouvement correcte. La représentation du mouvement est celle que nous avons décrite précédemment : un mouvement rigide 3D est un vecteur réel de dimension 6.

Dans ces conditions, nous allons développer un filtre de Kalman itératif dans le domaine «spatial» (et non temporel) ; lors de l'itération $K+1$, on dispose en entrée du filtre de l'estimation du mouvement obtenue à partir des K premiers couples de points. Pour chaque nouveau couple, on itère le filtre. Par suite, on obtient une estimation du mouvement, notée \hat{p}_N^1 qui sera alors prise en compte comme estimation initiale lorsqu'on appliquera une deuxième fois le filtre, et ainsi de suite jusqu'à convergence de celui-ci.

- Vecteur de mesure de dimension 6 : $\chi' = \begin{bmatrix} M \\ M' \end{bmatrix}$ où M et M' sont des points 3D appariés, liés par la relation $M' = RM+t$.
- Vecteur d'état de dimension 6 : $p = \begin{bmatrix} r \\ t \end{bmatrix}$, d'après la représentation du mouvement choisie.
- Equation d'état : $p_i = p_{i-1} = p$ (le mouvement appliqué sur la première vue est supposé être le même pour tous les points de cette vue).

8 La représentation du mouvement

8.1 Le mouvement rigide

Un mouvement appliqué à un objet indéformable est dit *rigide* [42] lorsque les relations géométriques (par exemple, la distance entre deux points) dans cet objet sont invariantes par application du mouvement. Tout mouvement rigide se décompose de manière unique en une translation et une rotation autour d'un axe passant par l'origine du repère.

Soit m le mouvement rigide et P un point de l'espace (on suppose que l'on applique la rotation puis la translation), on a :

$$m: P \rightarrow m(P) = P' \\ \text{avec } P' = RP + t$$

avec t , un vecteur de dimension 3 : le vecteur de translation et R , une matrice 3x3 : la matrice de rotation, orthogonale et de déterminant égal à 1.

On dira que (R, t) représente le *mouvement rigide*.

Par suite, 12 paramètres sont nécessaires pour définir un mouvement rigide ; toutefois, on peut montrer qu'un mouvement rigide 3D a seulement 6 degrés de libertés (3 pour la translation et 3 pour la rotation, les 9 paramètres de la rotation n'étant pas indépendants entre eux).

8.2 La représentation choisie

Nous avons choisi de représenter la rotation par un vecteur de dimension 3, $r = [a \ b \ c]^t$, dont la direction définit l'axe de rotation et dont la norme est égale à l'angle de rotation.

Par convention, on notera par la suite \tilde{r} la matrice anti-symétrique définie par :

$$\tilde{r} = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix} \quad (\text{EQ 66})$$

Dans ces conditions, la formule de Rodrigues permet d'établir le lien entre la matrice de rotation R et le vecteur r :

$$R = e^{\tilde{r}} = I_3 + f(\theta)\tilde{r} + g(\theta)\tilde{r}^2 \quad (\text{EQ 67})$$

où $\theta = \sqrt{a^2 + b^2 + c^2}$ est l'angle de rotation,

Algorithme du Filtre de Kalman Etendu

- Prédiction de l'état :

$$\hat{p}_{i|i-1} = h_i(\hat{p}_{i-1})$$

- Prédiction de la matrice de covariance de l'état :

$$S_{i|i-1} = \frac{\partial h_i}{\partial p_i} \cdot S_{i-1} \cdot \frac{\partial h_i^T}{\partial p_i} + Q_{i-1}$$

- Matrice de gain :

$$K_i = S_{i|i-1} A_i^T (A_i S_{i|i-1} S_i^T + \Lambda_{\xi_i})^{-1}$$

- Mise à jour de l'estimation de l'état :

$$\hat{p}_i = \hat{p}_{i|i-1} + K_i (y_i - A_i \cdot \hat{p}_{i|i-1}) = \hat{p}_{i|i-1} - K_i f_i(x_i, \hat{p}_{i|i-1})$$

- Mise à jour de la matrice de covariance de l'état :

$$S_i = (I - K_i \cdot A_i) S_{i|i-1}$$

- Initialisation :

$$S_{0|0} = \Lambda_{p_0}$$

$$\hat{p}_{0|0} = E[p_0]$$

7.2.2. Le Filtre de Kalman Étendu

Dans le cas où les équations du système ne sont pas linéaires, on parlera de filtre de Kalman étendu. Le principe de la mise en place des équations du filtre est alors de linéariser de manière itérative selon un développement de Taylor l'équation du système autour de l'estimée de l'état précédente.

Par suite, on obtient en linéarisant $f_i(x_i', p_i)$ autour du point $(x_i, \hat{p}_{i|i-1})$:

$$\begin{aligned} f_i(x_i', p_i) = & f_i(x_i, \hat{p}_{i|i-1}) + \frac{\partial f_i(x_i, \hat{p}_{i|i-1})}{\partial x_i'}(x_i' - x_i) + \frac{\partial f_i(x_i, \hat{p}_{i|i-1})}{\partial p_i}(p_i - \hat{p}_{i|i-1}) \\ & + O((x_i' - x_i)^2) + O((p_i - \hat{p}_{i|i-1})^2) \end{aligned} \quad (\text{EQ 63})$$

Ce développement peut alors s'écrire :

$$y_i = A_i p_i + \xi_i \quad (\text{EQ 64})$$

en posant :

$$A_i = \frac{\partial f_i(x_i, \hat{p}_{i|i-1})}{\partial p_i} \quad (\text{EQ 65})$$

$$y_i = -f_i(x_i, \hat{p}_{i|i-1}) + \frac{\partial f_i(x_i, \hat{p}_{i|i-1})}{\partial p_i} \hat{p}_{i|i-1}$$

$$\xi_i = \frac{\partial f_i(x_i, \hat{p}_{i|i-1})}{\partial x_i'}(x_i' - x_i) = -\frac{\partial f_i(x_i, \hat{p}_{i|i-1})}{\partial x_i'} e_i$$

$$E[\xi_i] = 0$$

$$E[\xi_i \xi_i^T] = \frac{\partial f_i(x_i, \hat{p}_{i|i-1})}{\partial x_i'} R_i \frac{\partial f_i(x_i, \hat{p}_{i|i-1})^T}{\partial x_i'} = \Lambda_{\xi_i}$$

L'algorithme du filtre de Kalman étendu est alors donné par les équations ci-dessous.

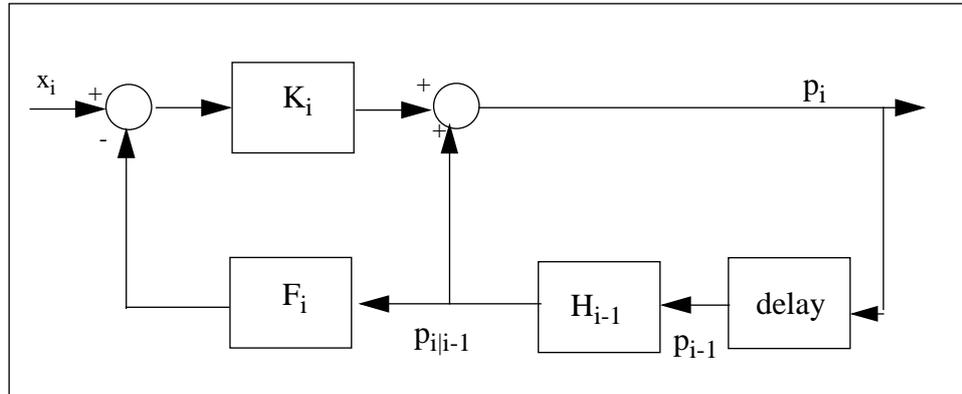


Figure 11. Filtre de Kalman : Bloc Diagramme

Le filtre de Kalman est, dans le cas général, un estimateur sans biais et de variance minimale. De plus, si les distributions probabilistes modélisant le système sont gaussiennes, alors il s'agit d'un estimateur selon le maximum de vraisemblance. En fait, le filtre de Kalman peut être vu comme la minimisation de la somme des distances de Mahalanobis entre le paramètre à estimer et l'ensemble des informations disponibles sur ce paramètre ; avec les notations précédentes, si p_0 est l'estimation initiale du paramètre p connue a priori, de matrice de covariance S_0 , alors le filtre de Kalman fournit le vecteur p qui minimise le critère suivant :

$$(p - p_0)^T S_0^{-1} (p - p_0) + \sum_i (x_i - F_i p_i)^T R_i^{-1} (x_i - F_i p_i) \quad (\text{EQ 61})$$

En d'autres termes, si \hat{p}_i est l'estimation optimale de p à l'instant t_i fournie par le filtre de Kalman, c'est aussi la solution de la fonction de coût suivante, où M est une matrice définie positive arbitraire :

$$C(\hat{p}_i) = E[(\hat{p}_i - p_i)^T M (\hat{p}_i - p_i)] \quad (\text{EQ 62})$$

Cette équation montre que le filtre de Kalman prend explicitement en compte le bruit de mesure et le pondère ; plus le bruit sur la i -ème mesure est grand, plus R_i^{-1} est petite et, par conséquent, moins le i -ème terme contribue à l'estimée finale.

Finalement, le filtre de Kalman est un algorithme permettant d'estimer de façon optimale l'état d'un système dynamique linéaire à partir d'un ensemble d'informations bruitées.

la matrice de gain K_i et mettre à jour, à l'instant t_i , l'estimation \hat{p}_i de l'état et sa matrice de covariance S_i .

Algorithme du Filtre de Kalman

- Prédiction de l'état :

$$\hat{p}_{i|i-1} = H_{i-1} \cdot \hat{p}_{i-1}$$

- Prédiction de la matrice de covariance de l'état :

$$S_{i|i-1} = H_{i-1} \cdot S_{i-1} \cdot H_{i-1}^T + Q_{i-1}$$

- Matrice de gain :

$$K_i = S_{i|i-1} F_i^T (F_i S_{i|i-1} F_i^T + R_i)^{-1}$$

- Mise à jour de l'estimation de l'état :

$$\hat{p}_i = \hat{p}_{i|i-1} + K_i (x_i - F_i \cdot \hat{p}_{i|i-1})$$

- Mise à jour de la matrice de covariance de l'état :

$$S_i = (I - K_i \cdot F_i) S_{i|i-1}$$

- Initialisation :

$$S_{0|0} = \Lambda_{p_0}$$

$$\hat{p}_{0|0} = E[p_0]$$

naissance du modèle d'évolution du processus. Le système est représenté dans l'espace d'état par l'équation d'état et par l'équation d'observation (observation de l'état par l'intermédiaire de la mesure).

Si l'on note p le vecteur d'état et x le vecteur de mesure, alors un système dynamique à temps discret est décrit par les équations suivantes :

$$\left\{ \begin{array}{l} p_{i+1} = h_i(p_i) + n_i, i \in \{0, 1, \dots\} \\ f_i(x_i', p_i) = 0 \end{array} \right. \quad (\text{EQ 58})$$

où n_i est un vecteur aléatoire modélisant un bruit blanc sur le système, de moyenne nulle et de matrice de covariance Q_i :

$$\left\{ \begin{array}{l} E[n_i] = 0 \\ E[n_i \cdot n_i^T] = Q_i \end{array} \right. \quad (\text{EQ 59})$$

En outre, les mesures sont elles-mêmes perturbées par un bruit blanc e_i de moyenne nulle et de matrice de covariance R_i , x_i étant la mesure réelle observée :

$$x_i = \mathbf{x}_i' + e_i$$

$$\text{avec : } \left\{ \begin{array}{l} E[e_i] = 0 \\ E[e_i \cdot e_j^T] = \delta_{ij} \cdot R_i \end{array} \right. \quad \text{où } \delta_{ij} \text{ est le symbole de Kronecker.}$$

On suppose que ces bruits sont indépendants, c'est-à-dire : $E[e_i \cdot n_j^T] = 0$, pour tout i et j .

7.2.1. Le filtre de Kalman

Le système peut être décrit par l'intermédiaire d'équations linéaires :

$$\left\{ \begin{array}{l} p_{i+1} = H_i p_i + n_i, i \in \{0, 1, \dots\} \\ x_i = F_i p_i + e_i \end{array} \right. \quad (\text{EQ 60})$$

Dans ces conditions, il est possible d'appliquer l'algorithme du filtre de Kalman ci-dessous qui permet d'obtenir à l'instant t_i la meilleure prédiction $\hat{p}_{i|i-1}$ de l'état p_i en fonction de l'estimation précédente de l'état \hat{p}_{i-1} ; on peut alors prédire la nouvelle matrice $S_{i|i-1}$ de covariance de l'état, calculer

de séquences d'images, la reconstruction de scènes 3D, la poursuite de cibles.

Dans ces perspectives, le filtrage de Kalman apparaît comme une méthode efficace pour estimer le déplacement à partir d'appariements de primitives (points, droites, segments) entre les différentes vues.

Dans [48], les auteurs traitent les problèmes de l'estimation des déplacements d'une caméra à partir de points et de droites appariés entre différentes vues, dans un environnement statique, et de la reconstruction de cet environnement. Ils montrent que l'appariement de points ou de droites entre plusieurs vues suivi par un filtre de Kalman ou une autre technique de moindres carrés pour estimer le mouvement est une méthode fiable et performante pour résoudre le problème même en présence de bruit dans les images.

Zhang et Faugeras utilisent dans [45] un filtre de Kalman qui apparaît comme la méthode la plus adaptée pour son efficacité et sa précision dans le cadre de l'estimation de mouvement 3D. Les primitives utilisées sont ici des segments de droites pour lesquels une nouvelle représentation est proposée, notamment quant à la modélisation de l'incertitude sur les paramètres du segment.

Dans [46], le problème de l'estimation de mouvement d'objets dans une séquence longue de triplets stéréoscopiques est formulé comme un problème de suivi auquel on peut appliquer immédiatement un filtre de Kalman étendu. Des segments 3D sont utilisés. Un modèle cinématique est construit pour chaque primitive étant entendu que chacune est attachée à un objet en mouvement. Dans ce contexte, le filtre de Kalman est utilisé pour *prédire* l'apparition de la primitive à l'instant suivant ; cette prédiction est alors utilisée pour *vérifier* si la primitive est effectivement présente dans l'image suivante et pour, éventuellement, la *recaler* avec la primitive réelle lui correspondant dans cette nouvelle image ; enfin, le recalage est utilisé pour la mise à jour du modèle cinématique. Par conséquent, mise en correspondance et estimation sont étroitement liés dans cette approche par l'intermédiaire du filtrage de Kalman.

7.2 Les équations générales du filtre de Kalman et du filtre de Kalman étendu

Le filtrage de Kalman est une approche récursive des techniques d'estimation optimales. Il fournit la meilleure estimée a posteriori d'une variable observée, lorsque celle-ci est une variable d'état d'un modèle dynamique connu et lorsque l'on connaît la statistique des bruits sur les mesures et le modèle.

Il permet alors de résoudre le problème de l'estimation pour des systèmes dynamiques en faisant appel à leur représentation d'état. Il requiert la con-

$$S(x, y) = [x, y, z(x, y)]^T, \quad (x, y) \in \mathfrak{R}^2$$

Pour calculer le déplacement à partir de ces appariements, une technique des moindres carrés est utilisée. Il s'agit donc de chercher la transformation (R, t) qui minimise la fonction suivante :

$$F(R, t) = \frac{1}{N} \cdot \sum_{i=1}^N \|RM_i + t - M_i'\|^2 \quad (\text{EQ 57})$$

où N est le nombre de paires (M_i, M_i') , avec M_i point de la première vue apparié avec M_i' dans la seconde vue.

Pour résoudre ce problème d'optimisation, Zhang utilise les représentations du quaternion et du quaternion dual. En ce qui nous concerne, nous résoudrons le problème de calcul du mouvement par un filtre de Kalman étendu, qui peut être vu comme une implantation récursive de la technique des moindres carrés pondérés.

7 Le filtrage de Kalman

Le filtre de Kalman joue un rôle essentiel en Automatique et Traitement du Signal. Dans un contexte d'estimation optimale et récursive, il permet notamment de résoudre des problèmes de mise en correspondance de données (étant donnés deux ensembles de mesures relatives à un ensemble de paramètres, comment appairer ces ensembles ?) ou de fusion de données (étant donnés deux ensembles de données de natures différentes, comment décider s'ils correspondent à un même ensemble de paramètres et, dans ce cas, comment les fusionner pour obtenir la «meilleure» estimation ?).

Un aspect du profil optimal du filtre de Kalman est l'utilisation de toutes les informations disponibles ; il traite l'ensemble des mesures en fonction de leur précision et de la connaissance que l'on en a, pour fournir l'estimée optimale d'un ensemble de paramètres.

7.1 Le filtre de Kalman et l'estimation de mouvement

L'analyse de mouvement tri-dimensionnel dans une séquence d'images occupe une place de choix dans le domaine de la vision par ordinateur et de la robotique, et apparaît comme un problème essentiel pour l'analyse et l'interprétation de scènes 3D. Avec le développement des systèmes de stéréovision, le calcul des déplacements à partir d'informations 3D est devenu un point d'intérêt tout spécialement attractif puisque ses applications touchent les domaines de la navigation autonome de robots mobiles, l'analyse

Chapitre 3 : Estimation de mouvement 3D par filtrage de Kalman

L'objectif de cette étude est de calculer de manière précise le mouvement entre deux cartes de profondeur successives afin de recalculer les deux vues. Notre algorithme s'inscrit dans celui de Zhang [22][23] dont la finalité est le recalage de deux nuages de points 3D dans le contexte de navigation autonome d'un véhicule planétaire utilisant la vision. Il concerne, en l'occurrence, exclusivement l'estimation des paramètres du mouvement 3D entre les deux vues et la comparaison avec les méthodes d'estimation de mouvement déjà développées, et non pas la mise en correspondance des points 3D. Pour plus de détails sur l'ensemble de la méthode de recalage de cartes de profondeur, on se reportera à [22].

Après avoir présenté succinctement l'algorithme de mise en correspondance de cartes de Zhang ainsi que la représentation du mouvement choisie, nous exposerons les équations générales du filtre de Kalman et son application dans le cadre de l'estimation de mouvement. Ceci nous conduira à développer un filtre de Kalman étendu pour résoudre notre problème. Nous serons alors en mesure de comparer les résultats de cette nouvelle méthode d'estimation de mouvement avec celle de Zhang.

6 Les algorithmes de mise en correspondance et d'estimation de mouvement

L'algorithme de recalage de Zhang essaie d'apparier les points 3D M_i de la première vue après leur avoir appliqué l'estimée du déplacement (R, t) , avec leurs plus proches voisins de la deuxième vue.

Les méthodes de recalage de cartes de profondeur peuvent être divisées en deux catégories [43][44] : si les deux cartes sont prises de deux points de vue très différents, ie. si le mouvement est grand, alors les approches basées sur des courbes ou des points critiques sont utilisées ; dans le cas où le mouvement est petit ou si l'on dispose d'une bonne estimation initiale du mouvement, l'approche basée sur l'ensemble des données est préférable. C'est cette seconde approche qui est utilisée par Zhang dans l'algorithme de recalage ; il utilise une technique qui s'appuie sur l'analyse statistique des distances d'appariements pour écarter les points erronés.

Désormais, une carte de profondeur est considérée comme une surface S de la forme :

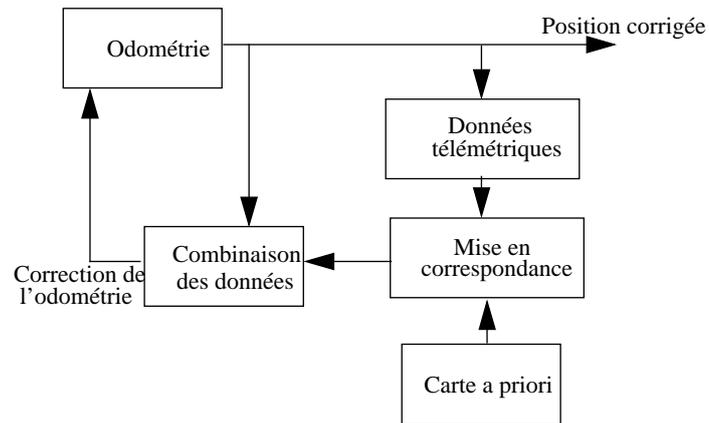


Figure 9. Le système de navigation du robot autonome Blanche

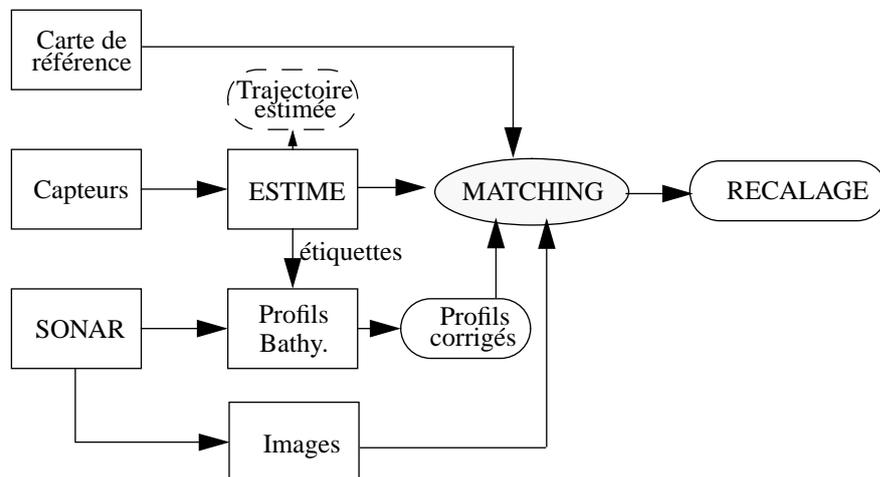


Figure 10. Synoptique du système de navigation référencée terrain

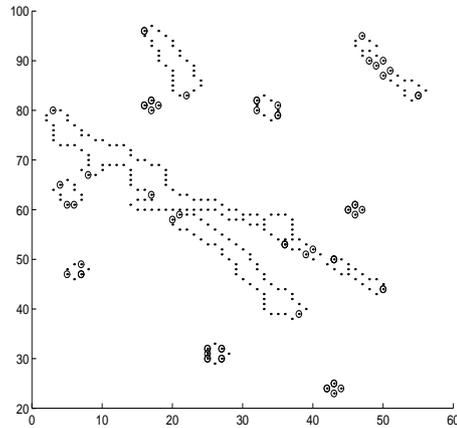


Figure 8. Les points critiques

5.2 Perspectives

Rappelons ce que nous souhaitons obtenir à terme : un algorithme de mise en correspondance destiné à la navigation autonome d'un véhicule sous-marin. Nous disposons d'une carte de référence embarquée (et éventuellement pré-traitée : segmentation en contours et extraction de points critiques) et d'une séquence d'images ou de profils bathymétriques. D'autre part, le véhicule et/ou le navire de surface sont équipés de différents capteurs d'estime : chaque image ou profil de la séquence est «étiqueté» par les données provenant de ces capteurs, qui, dans le cas où les fonds survolés sont peu chahutés (pouvoir discriminant faible), sont seuls à fournir une estimée exploitable de la position. Toutefois, ils ont l'inconvénient de dériver dans le temps ; à cet endroit, l'algorithme de mise en correspondance se devra alors de fournir une estimation de la position et donc de la trajectoire de l'engin (figure 9).

Une telle méthode se retrouve, par exemple, dans [41], dans un contexte de robotique terrestre, où un algorithme de mise en correspondance entre une carte de référence et des données télémétriques est utilisé pour pallier à la dérive d'un système de navigation inertiel : les deux approches conjuguées fournissent alors une estimation précise de la position (figure 10). De même, Bergem a observé dans [36] que, lorsque vitesse et mise en correspondance sont utilisées conjointement dans le système, les erreurs sont très nettement réduites et le recalage de la trajectoire meilleur.

- Recherche de points critiques :

A ce niveau, nous traitons des contours chaînés.

Préalablement à tout processus de mise en correspondance, nous nous sommes intéressés à l'extraction de points critiques qui pourraient faciliter la mise en correspondance. Les points de forte courbure apparaissent, dans ce cadre, des éléments critiques aisément identifiables sur nos contours. Nous avons envisagé deux méthodes géométriques de calcul de courbure [50] en intégrant un facteur de lissage k . En outre, nous souhaitons nous affranchir de méthodes différentielles introduisant un bruit non négligeable dans notre cas.

La courbure en un point C est définie par :

$$C = \frac{d\theta}{ds} \quad (\text{EQ 55})$$

où s est l'abscisse curviligne et θ l'angle de la tangente à la courbe.

Sur un maillage discret, la courbe ne présente une courbure non nulle que sur les points du maillage et la courbure ne peut prendre qu'un nombre restreint de valeurs, qui souvent ne sont pas significatives de l'allure globale de la courbe. Dans l'objectif de pallier à cet inconvénient, on définit une courbure discrète normalisée à l'ordre k , utilisant, non plus les points adjacents au point où l'on calcule la courbure, mais d'autres points sur la courbe :

$$C_k(P_i) = \frac{\text{angle}(P_{i-k}P_i, P_iP_{i+k})}{\|P_{i-k}P_i\| + \|P_iP_{i+k}\|} \quad (\text{EQ 56})$$

La figure suivante (figure 8) représente les contours et les points critiques obtenus après lissage à l'ordre 3 et seuillage de la courbure.

- Si $e[i, j]$ est la carte de profondeur initiale, avec $i \in [0, N - 1]$ et $j \in [0, N - 1]$, soit $c[i, j]$ la carte des contours.

$$e \otimes \nabla^2 G[i, j] = \sum_{m=-M}^{m=M} \sum_{n=-M}^{n=M} \nabla^2 G[m, n] \cdot e[i-m, i-n] = v[i, j] \text{ pour } i \in [0, N - 1] \text{ et } j \in [0, N - 1] \quad (\text{EQ 54})$$

- Détection des passages par zéro.

Sur la figure suivante (figure 7), on représente une carte de falaises obtenue pour une sous-image du Canyon du Var.

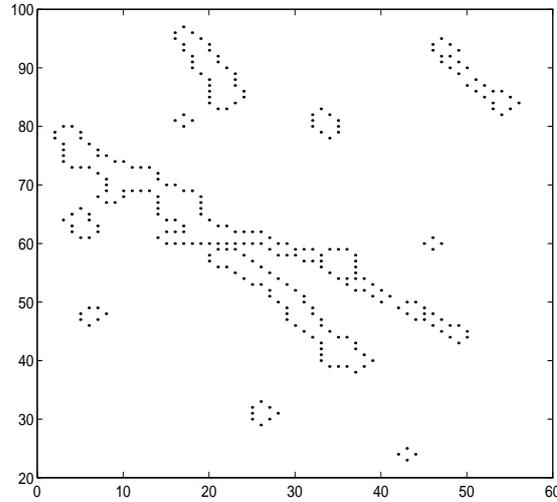


Figure 7. Carte de contours

Chacun des contours est chaîné puis représenté par un codage de Freeman à huit directions (chaque point d'un contour est codé par la direction suivant laquelle on obtient son plus proche voisin en 8-connexité). Le codage de Freeman peut donc être vu comme un codage de la tangente en tout point du contour ; dans ce cadre, toute notion de calcul de courbure discrète (et par suite d'extraction de points critiques) est alors envisageable.

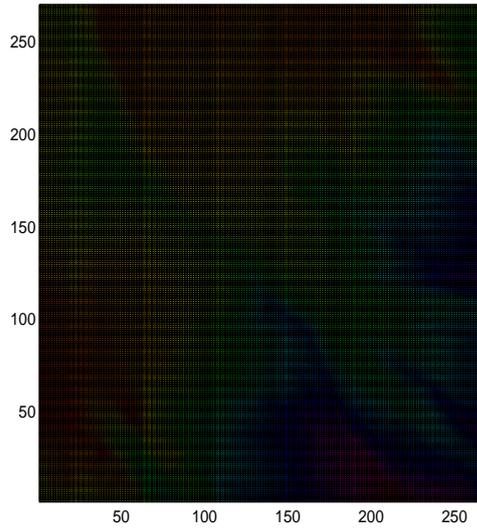


Figure 6. Niveaux de gris de la carte bathymétrique du canyon du Var

Nous nous sommes intéressés à l'approche de Rodriguez [17] essentiellement quant au type de données traitées. Les données sur lesquelles nous travaillons (bathymétrie du canyon du Var, (figure 6)) sont des cartes de profondeur.

Convertir ces données en cartes de falaise est non seulement un moyen de réduire la taille des données traitées mais aussi il s'agit là d'une méthode préalable de *segmentation* en contours fermés qui ouvre la voie à toute méthode de mise en correspondance, que celle-ci soit faite par l'intermédiaire d'un codage de Freeman ou par une corrélation directe, par exemple.

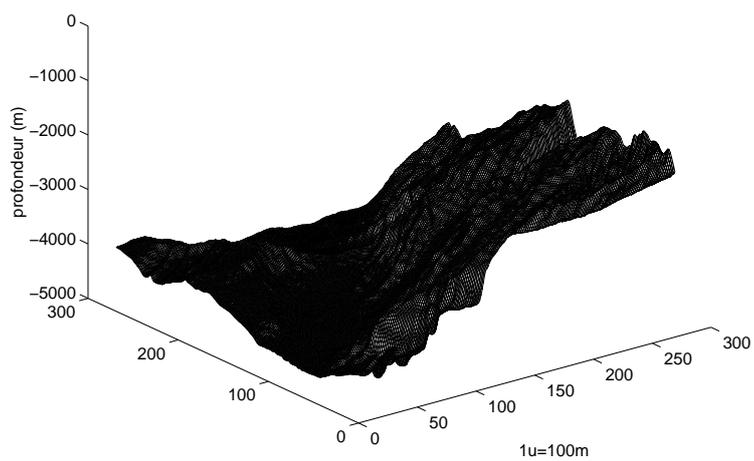
Rappelons les étapes principales de cet algorithme :

- Convolution de la carte de profondeur avec un filtre Laplacien d'une Gaussienne :

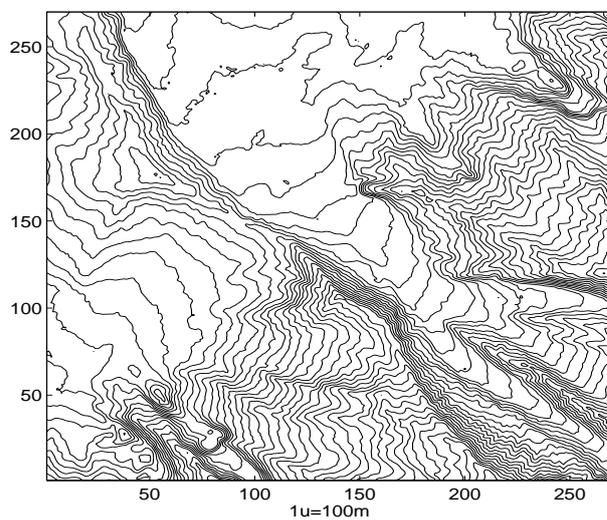
$$\nabla^2 G[i, j] = \frac{1}{2\pi\sigma^6} (r^2 - 2\sigma^2) \exp\left(\frac{-r^2}{2\sigma^2}\right) \text{ pour } i \in [-M, M] \text{ et } j \in [-M, M] \quad (\text{EQ 53})$$

avec $r^2 = i^2 + j^2$ et M , la taille du filtre discret.

5.1 Représentation et traitement des données



(a)



(b)

Figure 5. Le canyon du Var (EM12) : Visualisation du relief (a) et isobathes (b)

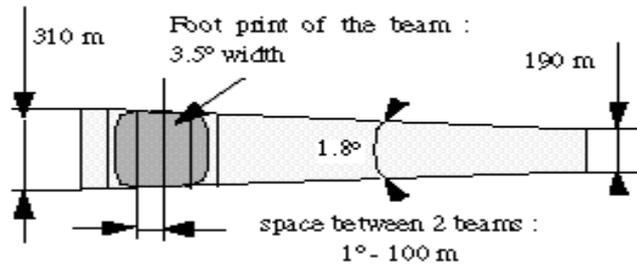


Figure 4. Intersection entre émission et réception

- Comparaison EM12 - EM1000 (tableau 1)

Une comparaison est possible uniquement pour des fonds dont la profondeur n'excède pas 1000m (limite d'utilisation de l'EM1000).

	Profondeur	Portée	Résolution Longitudinale	Résolution Transversale	
				Vert.	Obliquité Moy.
EM12	1000 m	1000 m	15 à 20 m	25 m	20 m
EM1000	800 m	880 m	42 à 48 m	42 m	23 m

Tableau 1. Caractéristiques moyennes EM12-EM1000

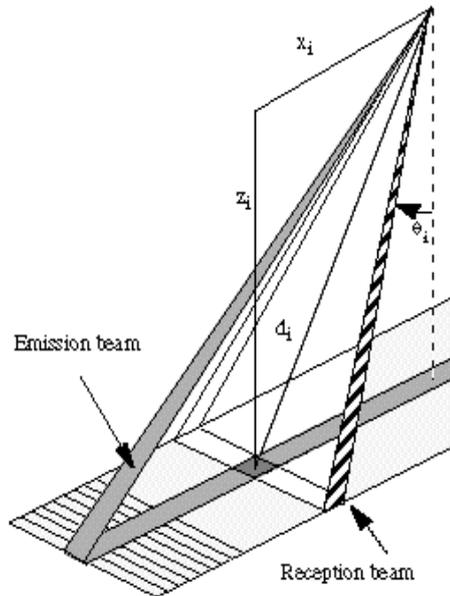


Figure 3. Méthode d'acquisition

- Principe de fonctionnement du système sondeur multi-faisceaux [40]

Les exemples numériques sont ceux de l'EM12 dual pour des fonds de 6000m.

Le sondeur utilise la technique des faisceaux croisés où la zone insonifiée, transversale au navire, est longue et étroite (1.8 degrés) alors qu'à la réception l'angle est très ouvert longitudinalement (20 degrés), ce qui permet d'éviter un asservissement en tangage et de compenser les mouvements de lacet. A la réception les faisceaux sont donc stabilisés seulement en roulis, alors qu'à l'émission, chaque secteur est stabilisé en roulis et en tangage pour gagner en niveau de signal et assurer une meilleure régularité des sondages.

Le faisceau à l'émission est de 150° dans la direction transversale ou latérale (*across direction*) et de 1.8° dans la direction longitudinale (*along direction*) (figure 3). A la réception, on dispose de 162 faisceaux espacés de 1° environ ; la largeur de ces faisceaux est de 3.5° dans la direction transversale et 20° dans la direction longitudinale (figure 4).

- Le sondeur EM-12 [39]

L'EM12 est un système multi-faisceaux simple ou dual (lorsque deux sondeurs identiques symétriques synchronisés sont placés à babord et tribord du navire) fonctionnant par des fonds de 100 à 10000 mètres. Le nombre de faisceaux est de 81 pour le système simple (respectivement 162 pour le système dual), espacés de 1.125 degrés, ce qui donne une couverture totale de 90 degrés (respectivement 150 degrés). Les données dont nous disposons ont été obtenues à partir du système dual.

L'EM12 possède deux modes de fonctionnement :

- ⇒ Le mode profond («DEEP»), pour des fonds de plus de 500 mètres. Un cycle est composé de 5 émissions successives de 10 millisecondes sur trois fréquences d'émission différentes ; chaque émission correspond à un secteur d'insonification dans le plan transversal au navire.
- ⇒ Par faibles fonds, en mode secteur ou «SHALLOW», une seule émission de 2 millisecondes couvre la totalité du plan transversal.

- Le sondeur EM-1000.

Le nombre de faisceaux, leur espacement, varient en fonction de la profondeur, selon trois modes :

- ⇒ Petits fonds («Shallow») : 60 faisceaux espacés de 2.5 degrés, soit un secteur de 150 degrés. La couverture peut donc atteindre 7.4 fois la profondeur.
- ⇒ Moyen («Medium») : 48 faisceaux espacés de 2.5 degrés, soit un secteur de 120 degrés ; le couloir couvert est de 3.4 fois la profondeur.
- ⇒ Profond («Deep») : 48 faisceaux espacés de 1.25 degrés, soit un secteur de 60 degrés ; le couloir couvert est de 1.15 fois la profondeur.

Chapitre 2 : Quelques résultats appliqués à nos données.

Dans ce chapitre, nous présentons, dans un premier temps, quelques caractéristiques techniques concernant les systèmes multi-faisceaux EM12 et EM1000, puis nous exposerons des résultats essentiellement fondés sur l'algorithme de recalage de Rodriguez appliqués à nos propres données bathymétriques.

5 Les caractéristiques des sondeurs multi-faisceaux EM12 et EM1000

Les sondeurs EM12 et EM1000 sont des systèmes multi-faisceaux (figure 2). Ils sont dotés à la fois de systèmes de bathymétrie - à partir de laquelle on réalise des cartes d'isobathes, et d'imagerie - qui traduit la variabilité de la réponse acoustique du fond. En ce qui nous concerne, nous nous intéressons aux caractéristiques en mode bathymétrique.

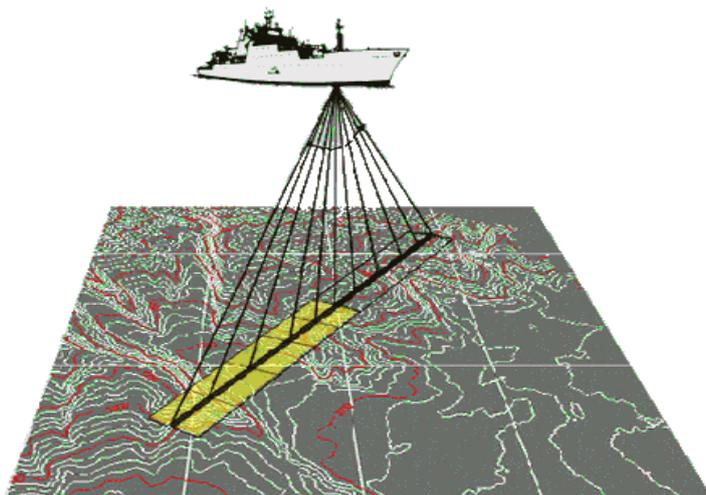


Figure 2. Le Sondeur EM12 dual multi-faisceaux du navire l'Atalante

$$f_N(x_1^k, x_2^k) = \frac{f_V(x_1^k, x_2^k)}{\iint_V f_N(x_1^k, x_2^k)(dx_1^k dx_2^k)} \quad (\text{EQ 52})$$

Les résultats expérimentaux montrent que les erreurs de positionnement sont essentiellement dues à l'absence de variations du fond, mais le système parvient toujours à se recalculer sur la trajectoire réelle. On observe l'absence de dérive par opposition aux systèmes de localisation utilisant les données cap et vitesse de l'engin.

4.3 L'approche de Cristi

Dans ses généralités, cette approche [38] est sensiblement équivalente à celle que nous avons décrit ci-dessus.

Il s'agit d'un algorithme de navigation d'un véhicule sous-marin dans un environnement connu a priori mais pouvant comporter éventuellement des obstacles non répertoriés ; il utilise les données sonar et une carte locale de l'environnement conjointement à un modèle dynamique de l'engin pour estimer, grâce à un filtre de Kalman, position et accélération du véhicule.

L'aspect original de l'algorithme réside dans l'utilisation de «fonctions de potentiel» pour définir la carte locale : elles permettent de définir les «régions d'attraction» correspondant aux obstacles et donc de modéliser l'environnement.

Ce système a été testé dans un bassin (de dimensions connues) mais des obstacles inconnus ont été introduits. Pour chacun des tests présentés, les données utilisées proviennent de cinq balayages successifs de 360° du sonar. Il est intéressant de noter que cette méthode est robuste en présence d'obstacles inconnus a priori et permet de les mettre en évidence.

4.2 L'approche de Bergem

Dans l'optique de la navigation autonome des engins sous-marins, Bergem [36][37] propose une méthode de navigation basée sur la reconnaissance des fonds avec un sonar multi-faisceaux. Le système retourne la position absolue de l'AUV (*Autonomous Underwater Vehicle*) en s'affranchissant de l'utilisation de balises et sans dérive dans le temps.

L'idée de base est la mesure des profondeurs sous différents angles à partir du sonar ; on dispose donc d'un profil précis du fond et la position absolue pourra être déterminée par mise en correspondance de ce profil avec une carte détaillée de la zone en question. Il est nécessaire qu'il y ait un minimum de variations en profondeur. D'autre part, quant à la précision des mesures, il faut trouver un compromis entre leur résolution, celle de la carte *a priori* et la taille du profil provenant du sonar. En fait, il existe une «distance critique» entre l'AUV et le fond qui dépend de la résolution de la carte de profondeur et des variations du fond.

Le principe de la méthode est basé sur un filtre de Kalman linéaire à temps discret où les mesures sont calculées à partir de la mise en correspondance et où le vecteur d'état est donné par (coordonnées cartésiennes et vitesse) :

$$X(k) = [x(k) \dot{x}(k) \quad y(k) \dot{y}(k)]^T$$

Un point délicat est l'estimation de la matrice de covariance des mesures car celles-ci proviennent de la mise en correspondance entre des données et un modèle (ie. la carte de référence). Pour cette estimation dynamique, on considère que la covariance des mesures est une fonction de la «qualité» de la mise en correspondance (la variance est d'autant plus petite que la mise en correspondance est bonne).

Pour cela, on définit une «fonction de coût» (*matching strength function*) $f_V(x, y)$ qui retourne un nombre positif décrivant la qualité de la mise en correspondance pour toute position (x, y) dans la zone V . Bergem montre alors que la matrice de covariance peut être estimée par :

$$R(k) = \begin{bmatrix} \sigma_{20} & \sigma_{11} \\ \sigma_{11} & \sigma_{02} \end{bmatrix} \quad (\text{EQ 50})$$

avec :

$$\sigma_{pq} = \iint_V f_N(x_1^k, x_2^k) (x_1^k - \bar{x}_1^k)^P (x_2^k - \bar{x}_2^k)^q dx_1^k dx_2^k \quad (\text{EQ 51})$$

portant plusieurs mouvements différents, les résultats - présentés sous forme de séquence d'images compensées par le mouvement estimé - sont des plus probants.

4 Vers la navigation référencée terrain...

Les véhicules autonomes terrestres jouissent d'avantages considérables : possibilité d'utiliser la vision optique, scènes composées d'objets aux structures géométriques bien définies... Au contraire, le monde sous-marin est sujet à de nombreux problèmes comme la présence des courants, de luminosité insuffisante, de formes difficilement modélisables géométriquement.

4.1 La vision pour le contrôle des véhicules sous-marins

L'aspect local et précis de la vision en fait une technique particulièrement attrayante dans le vaste domaine du contrôle des engins sous-marins. Les mesures relatives obtenues par les capteurs de vision peuvent être utilisées pour affiner le contrôle. En effet, dans le domaine de la navigation des engins sous-marins, il n'existe pas de systèmes fournissant des mesures de position globale ; les systèmes basés par exemple sur l'utilisation de balises acoustiques, ne peuvent couvrir que des zones limitées. En complément de ces systèmes, il est possible d'utiliser des capteurs locaux dont les mesures relatives fournissent directement une information sur les paramètres intéressants pour l'application en cours. Les auteurs [35] montrent que les images optiques sous-marines contiennent des caractéristiques opportunes par comparaison avec l'imagerie sonar : l'introduction de lumière artificielle affecte très peu l'environnement (dans le cas de capteurs sonar, l'environnement peut être modifié par l'addition de nouvelles sources d'énergie) ; la résolution est meilleure (en gardant à l'esprit que l'on se place dans le cas de capteurs locaux) ; le nombre de données est très grand. Pour pallier à certains problèmes dus au milieu sous-marin, l'approche combine un filtrage par convolution avec un filtre LoG (*Laplacian-of-Gaussian*) suivi par une corrélation binaire sur les images du signe du filtre LoG.

Le filtre Gaussien peut être vu comme un passe-bande qui permet, en ajustant sa variance, d'éliminer certains bruits dans l'image. D'autre part, la binarisation (image du signe de la convolution) réduit les effets de non-uniformité de luminosité ou de contrastes.

Ces techniques de vision ont été adaptées pour différentes applications sous-marines comme le suivi de cibles en mouvement, la création de mosaïques ou encore la navigation.

où

$A^t=(a_1, \dots, a_n)$ est le vecteur de paramètres du mouvement,
 $p_i=(x_i, y_i)$ est un point de l'image,
 V_A est le vecteur flot optique,
 B est une matrice dans la forme dépend du modèle de mouvement
choisi mais dont les coefficients ne dépendent que des coordonnées de
 p_i .

Pour décrire la plupart des mouvements (rotation, translation, déformation, changement d'échelle), on choisira le modèle à 6 paramètres :

$$\begin{cases} u_A(p_i) = a_1 + a_2 \cdot x_i + a_3 \cdot y_i \\ v_A(p_i) = a_4 + a_5 \cdot x_i + a_6 \cdot y_i \end{cases} \quad (\text{EQ 47})$$

avec :

$$B(p_i) = \begin{bmatrix} 1 & x_i & y_i & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_i & y_i \end{bmatrix} \quad (\text{EQ 48})$$

L'image est décomposée en pyramides gaussiennes passe-bas.

L'estimation du paramètre $\Theta=(A^t, \xi)$ se fait de manière incrémentale dans l'approche du niveau le plus grossier L au plus fin O .

Pour résoudre le problème de l'estimation (les données sont bruitées, les modèles de mouvement ne sont que des approximations, et il y a toujours une contrainte de temps de calcul à respecter), les auteurs utilisent l'approche des estimateurs robustes et ramènent le problème à celui des moindres carrés pondérés :

$$\sum_{p_i \in R} \rho(r_i) = \sum_{p_i \in R} \frac{1}{2} \cdot w_i \cdot r_i^2 \quad (\text{EQ 49})$$

avec : $r_i = o_i - M(\Theta, p_i)$, les o_i étant les observations, Θ le paramètre à estimer, M le modèle.

$w_i = \frac{\psi(r_i)}{r_i}$ sont les poids définis à partir de la dérivée ψ de ρ

Cet algorithme a été appliqué à des images sous-marines en utilisant plusieurs modèles de mouvement. Les images étant fortement bruitées et com-

Si on choisit $\rho(t) = \frac{1}{2} \cdot t^2$, alors on retrouve les équations (36) et (37).

Quelques fonctions ρ possibles ...

$$\rho(t) = |t| \qquad \Psi(t) = \text{sgn}(t) \qquad (\text{EQ 42})$$

$$\rho(t) = \log\left(1 + \frac{1}{2} \cdot t^2\right) \qquad \Psi(t) = \frac{t}{1 + \frac{1}{2} \cdot t^2} \qquad (\text{EQ 43})$$

Quelques fonctions de poids ψ ...

$$\psi(t) = \begin{cases} \sin\left(\frac{t}{c}\right) & \text{si } |t| < c\pi \\ 0 & \text{si } |t| > c\pi \end{cases} \quad (\text{Andrew's sine}) \qquad (\text{EQ 44})$$

$$\psi(t) = \begin{cases} t\left(1 - \frac{t^2}{c^2}\right)^2 & \text{si } |t| < c \\ 0 & \text{si } |t| > c \end{cases} \quad (\text{Tukey's biweight}) \qquad (\text{EQ 45})$$

3.5 Une approche d'estimation robuste multi-échelles de mouvement

La méthode présentée ici [33][34] permet d'estimer de manière fiable et précise des modèles paramétrés 2D de mouvement ; elle permet de gérer la présence de plusieurs mouvements dans une image. Elle est fondée sur l'estimation robuste (M-estimateurs) dans un schéma multi-résolution (pyramides gaussiennes de Burt).

D'une manière générale, les modèles de mouvement 2D peuvent être mis sous la forme matricielle suivante :

$$V_A(p_i) = \begin{bmatrix} u_A(p_i) \\ v_A(p_i) \end{bmatrix} = B(p_i)A \qquad (\text{EQ 46})$$

raisonnement des estimateurs selon le maximum de vraisemblance (moindres carrés).

Le problème reste le même : on dispose de N données (x_i, y_i) , $i \in [1, N]$, il s'agit d'adapter un modèle à cet ensemble en ajustant M paramètres a_j , $j \in [1, M]$:

$$y(x) = y(x; a_1, \dots, a_M) \quad (\text{EQ 36})$$

Lorsque les erreurs sur les mesures peuvent être modélisées par une loi normale, le problème revient à minimiser la quantité ci-dessous par rapport aux paramètres a_j :

$$\sum_{i=1}^N \left(\frac{y_i - y(x; a_1, \dots, a_M)}{\sigma_i} \right)^2 \quad (\text{EQ 37})$$

Si l'on dérive cette expression par rapport à chacun des paramètres a_k , on obtient un système de M équations :

$$\sum_{i=1}^N \left(\frac{y_i - y(x_i)}{\sigma_i^2} \right) \frac{\partial}{\partial a_k} (y(x_i), \dots, a_k, \dots) = 0, \quad \forall k \in [1, M] \quad (\text{EQ 38})$$

Dans le cas où l'on ne connaît pas la loi de distribution de l'erreur sur les mesures, on peut généraliser l'équation (36) :

$$\sum_{i=1}^N \rho \left(\frac{y_i - y(x; a)}{\sigma_i} \right) \quad (\text{EQ 39})$$

Si l'on pose :

$$\psi(t) \equiv \frac{d}{dt} \rho(t) \quad (\text{EQ 40})$$

alors l'équation (37) se généralise aux cas des M-estimateurs :

$$\sum_{i=1}^N \frac{1}{\sigma_i} \psi \left(\frac{y_i - y(x_i)}{\sigma_i} \right) \cdot \frac{\partial}{\partial a_k} (y(x_i), a) = 0, \quad \forall k \in [1, M] \quad (\text{EQ 41})$$

On note $f(x, y, t)$ la fonction intensité, $p=(x, y)$ un point image et $d=(dx, dy)$ un déplacement dans le plan image. La relation entre déplacement et vitesse est, sous forme discrète :

$$\vec{d} = \Delta t \cdot \vec{V} \quad (\text{EQ 30})$$

où Δt est l'intervalle de temps entre deux images dans la séquence considérée.

On suppose qu'il y a invariance dans le temps de la luminance d'un point p qui se déplace de d entre t et $t+dt$, ce qui se traduit par l'équation de contrainte du mouvement apparent :

$$f(x + dx, y + dy, t + dt) = f(x, y, t) \quad (\text{EQ 31})$$

Un développement limité au premier ordre au point (x, y, t) donne :

$$f(x + dx, y + dy, t + dt) = f(x, y, t) + \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial t} dt + o^2 \quad (\text{EQ 32})$$

D'après les équations (31) et (32), on obtient :

$$\frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial t} dt = 0 \quad (\text{EQ 33})$$

Si l'on note $\vec{\nabla} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$ le gradient spatial de l'intensité, alors la relation (33) s'écrit sous forme vectorielle :

$$\vec{V} \cdot \vec{\nabla} f + \frac{\partial f}{\partial t} = 0 \quad (\text{EQ 34})$$

On en déduit la composante de la vitesse mesurable (observation locale) : elle est parallèle au gradient de l'intensité :

$$V^\perp = -\frac{\partial f}{\partial t} \frac{\vec{\nabla} f}{\|\vec{\nabla} f\|} \quad (\text{EQ 35})$$

3.4 L'estimation robuste

Par *robuste*, on entend l'insensibilité aux petites déviations à partir des hypothèses idéales pour lesquelles l'estimateur est optimal. Parmi tous les estimateurs robustes, on trouve la classe des **M-estimateurs** qui suivent le

mation initiale influence le temps d'exécution de l'algorithme qui peut converger vers un minimum global ou local selon la précision de cette estimation.

3.3 La notion de flot optique

Un des objectifs de l'analyse de mouvement est de retrouver le mouvement 3D à partir des informations dynamiques contenues dans la séquence d'images.

Une telle étude passe par l'analyse du mouvement apparent des objets dynamiques dans la séquence et elle comprend les aspects suivants :

- Détection des zones en mouvement : attribuer une étiquette à chaque point de l'image pour indiquer s'il est en mouvement ou pas,
- Segmentation de l'image en régions de mouvement apparent différent. Cette étape est étroitement liée à l'aspect estimation, puisqu'il est nécessaire de disposer de l'estimation du champ de vitesses pour obtenir une segmentation correcte en régions homogènes,
- Estimation du champ de vitesse. Les méthodes d'estimation sont la mise en correspondance et les approches différentielles dont fait partie le flot optique,
- Identification ou étiquetage du champ de vecteurs en fonctions des configurations cinématiques.

L'approche différentielle :

L'information de mouvement est directement liée aux observations disponibles dans une image et essentiellement aux valeurs d'intensité des pixels. Le problème est l'estimation des deux composantes du vecteur vitesse V .

$$V = \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} \quad (\text{EQ 29})$$

Pour distinguer l'objet de l'arrière plan, $f(i, j, k)$ est définie par :

$$f(i, j, k) = t(i, j, k)\lambda(i, j, k) + b(i, j, k)[1 - \lambda(i, j, k)] \quad (\text{EQ 27})$$

où $t(i, j, k)$ et $b(i, j, k)$ sont deux processus indépendants correspondant respectivement à l'objet et au fond. En outre, $\lambda(i, j, k)$ est une fonction indicatrice égale à 1 lorsque le point (i, j, k) est un point objet et 0 sinon.

En fait, les auteurs supposent que la séquence d'images a été segmentée préalablement (ie. λ est connue pour chaque image dans la séquence).

Dans ces conditions, il s'agit d'estimer le mouvement 2D entre $f(i, j, k)$ et $f(i, j, k+1)$, c'est-à-dire on a la relation suivante :

$$f(x, k+1) = f(C(k)x + s(k), k) \quad (\text{EQ 28})$$

où

$$x = \begin{bmatrix} i \\ j \end{bmatrix}$$

$$C(k) = \begin{bmatrix} c_{11}(k) & c_{12}(k) \\ c_{21}(k) & c_{22}(k) \end{bmatrix}$$

$$s(k) = \begin{bmatrix} s_1(k) \\ s_2(k) \end{bmatrix}$$

A cet endroit, le choix de la méthode pour estimer les paramètres du mouvement est motivé par la robustesse en présence de bruit ; d'où la préférence pour la méthode de mise en correspondance plutôt que pour la méthode différentielle (qui fait appel à des calculs de dérivées). Les auteurs sont alors ramenés à un problème d'optimisation.

Dans le cas où le mouvement 2D est linéaire, ses paramètres sont retrouvés de façon exacte ; dans le cas non-linéaire, cet algorithme fournit une bonne approximation linéaire. Sa robustesse est prouvée de manière qualitative lorsque des images réelles sont utilisées.

Toutefois, nous pouvons reprocher à cet algorithme de nécessiter d'une bonne segmentation préalable des images et de la connaissance du mouvement initial pour amorcer l'algorithme d'estimation ; on notera que cette esti-

3.2 Estimation du mouvement d'un objet 3D se déplaçant dans un environnement 3D

Pour étudier le mouvement d'un objet tri-dimensionnel, les auteurs [32] se ramènent à l'étude du mouvement 2D de la projection centrale de cet objet ; c'est à partir de l'analyse des paramètres du mouvement de 2D, que celui-ci soit linéaire ou non, qu'ils pourront alors estimer le mouvement 3D (mais cette étape finale n'est pas présentée dans ce papier).

L'algorithme est basé sur la mise en correspondance 2D des projections centrales de l'objet 3D qui sont considérées comme des primitives. Eventuellement, il peut y avoir plusieurs objets en mouvement.

- Modélisation du mouvement 2D :

Que celui-ci soit linéaire ou non, que l'on introduise en facteur d'échelle uniforme (par l'intermédiaire d'une matrice de déformation) dans la projection de l'objet ou non, les auteurs concluent à une représentation unique du mouvement 2D sous la forme d'un modèle linéaire affine :

$$r' = Cr + s \quad (\text{EQ 26})$$

avec :

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

$$s = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}$$

où $c_{11}, c_{12}, c_{21}, c_{22}, s_1, s_2$ sont les paramètres du mouvement (rotation et translation).

- Modélisation de la séquence d'images :

Soit $f(i, j, k)$ la séquence d'images, avec i et j les coordonnées d'un point de l'image, et k le temps.

tout l'accent sur le recalage de cartes proprement dit ; les méthodes pour estimer la transformation entre les deux images sont souvent très simples, d'autant plus qu'il s'agit, le plus souvent, de transformations dans un plan (ie. à trois paramètres). Néanmoins, les techniques d'estimation de mouvement - en aval du recalage - peuvent être plus raffinées voire intrinsèquement liées à la méthode de mise en correspondance elle-même, comme nous avons pu le constater à travers la méthode de Zhang [22].

3.1 Mise en correspondance et estimation de mouvement

D'une manière générale, tout algorithme d'estimation de mouvement peut être divisé en trois parties essentielles [31] :

- Segmentation des images et extraction de primitives caractéristiques,
- Mise en correspondance de ces primitives entre les vues,
- Estimation du mouvement en utilisant ces appariements.

Soit P un vecteur de paramètres représentant une primitive, et P' la primitive appariée à P dans la seconde vue. L'ensemble $S = \{(P, P')\}$ est issu des étapes de segmentation puis de mise en correspondance. Le problème posé est alors : *étant donné S , estimer le déplacement entre les vues*. Si T représente la transformation à déterminer, alors on peut écrire :

$$P' = f(P, T)$$

où la nature de la fonction f dépend des primitives choisies, de leur représentation et de la représentation de la transformation T . En outre, il peut s'avérer que les données soient bruitées et que les observations ne satisfassent pas les contraintes de rigidité du mouvement. Si n et n' sont respectivement le bruit sur P et sur P' , alors on peut quantifier l'erreur d'estimation du mouvement par :

$$\sum_{(P, P') \in S} \|\tilde{P}' - f(\tilde{P}, T)\|^2 \quad (\text{EQ 25})$$

où $\tilde{P} = P + n$ et $\tilde{P}' = P' + n'$

Le problème est alors posé en termes de minimisation de la somme ci-dessus : *étant donné l'ensemble S , calculer l'estimation optimale du déplacement qui minimise l'erreur quadratique*.

pour réaliser la mise en correspondance par des techniques de corrélation, ceci dans le but de réduire le coût de calcul considérable causé par la corrélation directe.

Les images sont représentées à différents niveaux de résolution de 0 (le plus fin) à L (le plus grossier). Si l'on se place à deux niveaux consécutifs, un pixel du niveau le plus grossier couvre quatre pixels du plus fin ; pour chaque création d'un niveau plus grossier, la bande passante est divisée par deux.

Les résultats de la méthode de corrélation pour la mise en correspondance sont meilleurs lorsque les basses fréquences sont filtrées. C'est pourquoi une pyramide passe-bande (ou Laplacienne) définie selon l'algorithme de Burt, est utilisée pour décrire les images.

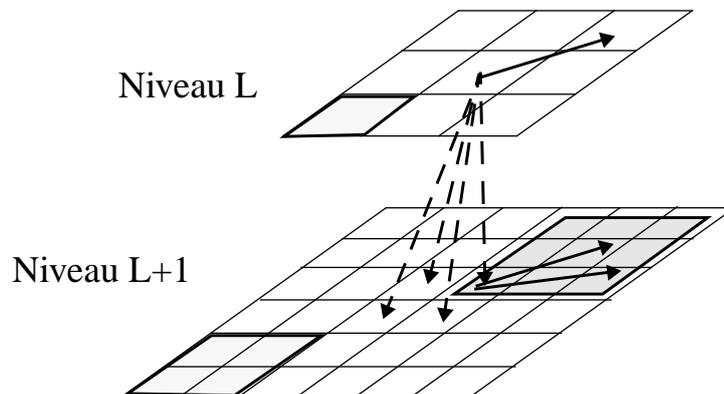


Figure 1. Estimation par approche hiérarchique

L'algorithme est alors le suivant (figure 1) : la mise en correspondance est faite initialement sur les images de fréquences les plus basses (niveau le plus grossier), ce qui fournit une estimation du déplacement utilisable au niveau inférieur (plus fin).

Les résultats montrent que cette approche hiérarchique est bien meilleure qu'une mise en correspondance par corrélation directe : le temps de calcul est très nettement réduit et il y a peu de faux appariements.

3 Analyse de séquences d'images

La limite entre mise en correspondance et estimation de mouvement dans une séquence d'images est floue : les techniques de mise en correspondance sont la plupart du temps utilisées à des fins d'estimation de mouvement entre deux vues. Les techniques évoquées dans le chapitre précédent mettent sur-

$$I(x, y, t) = \int \int_{\mathfrak{R}^2} I_0(x-u, y-v) g(u, v) dudv \quad (\text{EQ 20})$$

Extension aux signaux discrets :

Dans [28][29], Lindeberg montre qu'il est impossible de construire directement la représentation multi-échelles discrète en dimension 2 à partir d'un simple échantillonnage des équations établies dans le cas des signaux continus. Il montre que le passage au domaine discret doit s'effectuer en discrétisant l'équation de diffusion.

Soit f un signal discret en dimension 2 : $f : Z^2 \rightarrow \mathfrak{R}$, la représentation multi-échelles $L : Z^2 \times \mathfrak{R}_+ \rightarrow \mathfrak{R}$ doit satisfaire, pour $\gamma \in [0, 1]$ fixé, l'équation de diffusion discrète suivante :

$$\frac{\partial L}{\partial t} = \frac{1}{2} \cdot \nabla_\gamma^2 L = \frac{1}{2} ((1-\gamma) \nabla_1^2 L + \gamma \nabla_2^2 L) \quad (\text{EQ 21})$$

avec la condition initiale $L(\dots, 0) = f$ et :

$$\begin{cases} \nabla_1^2 L = f_{-1,0} + f_{1,0} + f_{0,-1} + f_{0,1} - 4 \cdot f_{0,0} \\ \nabla_2^2 L = \frac{1}{2} \cdot (f_{-1,-1} + f_{-1,1} + f_{1,-1} + f_{1,1} - 4 \cdot f_{0,0}) \end{cases} \quad (\text{EQ 22})$$

Dans le cas particulier où $\gamma = 0$, la famille de signaux est obtenue par :

$$L(x, y, t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} T(m, t) \cdot T(n, t) \cdot f(x-m, y-n) \quad (\text{EQ 23})$$

T étant appelé «analogue discret du noyau gaussien» :

$$T(n, t) = e^{-t} \cdot I_n(t) \quad (\text{EQ 24})$$

où $I_n(t)$ sont les fonctions de Bessel modifiées d'ordre entier.

2.6 Mise en correspondance et multi-résolution

L'algorithme présenté ici [30] ne met pas en correspondance des images de résolutions différentes, mais utilise la décomposition pyramidale d'une image afin d'adopter la stratégie hiérarchique «du plus grossier au plus fin»

- Principe de causalité : les structures aux niveaux grossiers doivent constituer des *simplifications* des structures correspondantes aux niveaux les plus fins,
- Isotropie et homogénéité du lissage : le lissage doit être invariant dans l'espace.

Ces deux critères conduisent à la conclusion suivante : la représentation espace- échelle doit satisfaire à l'équation de diffusion.

Si I_0 est l'image initiale et $I(x, y, t)$ est l'image à l'échelle t , alors $I(x, y, t)$ est solution de l'équation suivante :

$$\frac{\partial I}{\partial t} = \frac{1}{2} \nabla^2 I \quad (\text{EQ 18})$$

avec la condition initiale $I(x, y, 0) = I_0$.

Cas des signaux continus :

L'ensemble des signaux dérivés est construit par convolution du signal original continu avec une famille de noyaux gaussiens à un paramètre, le *paramètre d'échelle*.

Soit $I_0 : \mathfrak{R}^2 \rightarrow \mathfrak{R}$, le signal initial.

La représentation multi-échelles est alors définie par :

$$I : \mathfrak{R}^2 \times \mathfrak{R}_+ \rightarrow \mathfrak{R}$$

$$\left\{ \begin{array}{l} I(x, y, t) = g(x, y, t) \otimes I_0 \\ I(x, y, 0) = I_0 \end{array} \right. \quad (\text{EQ 19})$$

où $t = \sigma^2$ est le paramètre d'échelle,

et $g(x, y, t) = \frac{1}{2\pi t} \cdot e^{-(x^2 + y^2)/2t}$ le noyau gaussien.

La formulation du problème comme la résolution de l'équation de diffusion implique, dans le cas 2D, d'envisager de traiter un domaine infini ie. une plaque de surface infinie. Le calcul de la solution de l'équation (17) revient à évaluer l'intégrale de convolution :

$$G_k[I] = 4 \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) \cdot G_{k+1}\left(\frac{i+m}{2}, \frac{j+n}{2}\right) \quad (\text{EQ 16})$$

Dans ce formalisme, on a :

$$\nabla_k = G_k - \text{EXPAND}(G_{k+1}) \quad (\text{EQ 17})$$

La pyramide Laplacienne est une décomposition en bande passante du signal initial I . Elle constitue de plus un codage exact de I (la seule connaissance de $\nabla[I]$ permet de reconstruire I). Elle est utilisée non seulement pour des problèmes de compression de données mais aussi pour l'extraction de caractéristiques.

2.5.2. La représentation espace-échelle (*scale-space*)

Dans l'objectif de mettre en correspondance des cartes de résolutions différentes, la théorie de l'espace multi-échelles introduite par Witkin [27] fournit un cadre d'étude intéressant et élégant.

Il convient à cet endroit de préciser la différence entre Analyse Multi-Résolution et Analyse Multi-Echelles. Telle que nous l'avons décrite précédemment, l'analyse multi-résolution est une représentation pyramidale de l'image initiale : pour chaque niveau de la pyramide, on réalise un sous-échantillonnage suivi d'un lissage, la taille des images analysées (ie. une sous-image de la pyramide) est alors inférieure à la taille de l'image à analyser. En ce qui concerne l'analyse multi-échelles que nous décrivons ici, la taille du support de l'image reste constante quel que soit le niveau de résolution. Dans un processus d'extraction de caractéristiques ou de mise en correspondance, c'est l'analyse multi-échelles qui apparaît la plus appropriée : elle permet de filtrer pas à pas les informations inutiles à l'analyse pour conserver les plus pertinentes tout en gardant la possibilité de mettre aisément en correspondance des structures entre les échelles. Il s'agit en outre d'une représentation invariante par transformation euclidienne et qui préserve le même échantillonnage spatial quelle que soit l'échelle.

Telle qu'elle est présentée ici, la représentation multi-échelles d'un signal est vue comme un ensemble ordonné - obtenu par la variation d'un paramètre - de signaux dérivés du signal original à différentes échelles (ie. pour différentes valeurs du paramètre). La famille de signaux est obtenue par convolution de l'image originale par une Gaussienne (ou une dérivée) dont l'écart-type est le *paramètre d'échelle*.

Une telle représentation doit satisfaire à deux critères fondamentaux :

Burt définit l'opérateur *REDUCE* qui génère un niveau de la pyramide à partir de son prédécesseur.

$$G_k[I] = REDUCE(G_{k-1}[I]) \quad (\text{EQ 9})$$

$$G_k(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) \cdot G_{k-1}(2i + m, 2j + n) \quad (\text{EQ 10})$$

La fonction de poids (noyau du filtre) est choisie sous quatre contraintes. En les combinant, un noyau possible de taille 5 est :

$$\left(\frac{1}{4} - \frac{a}{2}, \frac{1}{4}, a, \frac{1}{4}, \frac{1}{4} - \frac{a}{2} \right) \quad (\text{EQ 11})$$

a étant une variable quelconque ; empiriquement, $a=0.4$ permet d'obtenir une fonction de lissage approximativement gaussienne.

La pyramide Laplacienne

La perte d'information lors du passage du niveau k au niveau $k+1$ peut être obtenue par la différence entre G_k et l'estimation de G_k à partir du niveau $k+1$, $E\{G_k\}$, avec :

$$E\{G_k\}(P) = [W \otimes (G_{k+1} \wedge 2)](P) \quad (\text{EQ 12})$$

avec

$\wedge 2$: opérateur de sur-échantillonnage permettant de ramener le niveau $k+1$ à la taille du niveau k ,

P : un point quelconque de l'image I .

Si on répète cette procédure à tous les niveaux de la pyramide, on obtient la pyramide Laplacienne :

$$\nabla_k(P) = G_k - E\{G_k\}(P) \quad k \in [0, L-1] \quad (\text{EQ 13})$$

$$\nabla_L \equiv G_L \quad (\text{EQ 14})$$

Burt parle de l'opérateur *EXPAND*, opérateur d'interpolation inverse de *REDUCE* :

$$G_k[I] = EXPAND(G_{k+1}[I]) \quad (\text{EQ 15})$$

résolution classique est le modèle pyramidal. Il permet de combiner parallélisme et récursivité.

Les algorithmes permettant de décomposer une image sous forme de pyramides [25] doivent posséder trois qualités essentielles :

- ne pas créer de structures non existantes dans l'image d'origine,
- traiter les structures présentes dans l'image indépendamment de leur orientation,
- ne pas alourdir la charge de calcul au détriment du bénéfice éventuel de cette nouvelle représentation (!).

La pyramide gaussienne

Les deux premiers points sont satisfaits par le choix d'un noyau de filtrage ou de convolution de type gaussien ; on parle alors de pyramide gaussienne ou filtre passe-bas multi-résolution [26].

Chaque résolution de l'image de départ est associée à un niveau de la pyramide gaussienne.

Celle-ci se construit par l'algorithme suivant :

$$\begin{aligned} G_0[I] &= I \\ G_k[I] &= (W \otimes G_{k-1}[I]) \vee 2, \quad k \in [1, L] \end{aligned} \quad (\text{EQ 7})$$

avec

W : filtre passe-bas,

\otimes : produit de convolution,

$\vee 2$: opérateur de sous-échantillonnage :

$$\text{taille}[G_k] = \frac{\text{taille}[G_{k-1}]}{2^d} \quad (\text{EQ 8})$$

d : dimension du signal (ici $d=2$)

I : image de départ,

L : dernier niveau de la pyramide (il correspond aux plus basses fréquences du signal I).

-
- Mettre à jour les appariements à l'aide d'une analyse statistique des distances,
 - Calcul du mouvement entre les deux vues en utilisant les appariements retenus lors de l'étape précédente,
 - Appliquer cette transformation à tous les points de la première vue.

L'intérêt de cet algorithme est qu'il ne fait appel à aucun pré-traitement ni segmentation préalables des images ; il est robuste de par l'analyse statistique des distances qui élimine entre deux itérations les appariements erronés ; il fournit une estimation précise du mouvement lorsque l'estimation initiale est correcte et il peut tenir compte des incertitudes sur les données si l'on utilise un filtre de Kalman comme méthode d'estimation du mouvement.

2.5 Vers une approche multi-échelles ...

Les cartes bathymétriques qu'il nous faut mettre en correspondance proviennent de deux sonars ayant des résolutions différentes : ce que l'on «voit» dans l'une des cartes n'est pas ce que l'on voit dans l'autre (fréquences différentes, résolutions et précisions différentes, couvertures spatiales différentes). Dans [24], on parlera de *focus d'attention* : «pour chaque entité de l'image, il existe une résolution donnée dans laquelle la recherche (ie. détection) et l'extraction (ie. localisation) sont facilitées». Des détails peu significatifs dans une image de haute résolution disparaissent pour une résolution réduite.

La représentation multi-résolution des images est un outil très utilisé en vision artificielle ; elle permet de décrire une image par une hiérarchie de structures. Elle est également un moyen privilégié pour réduire la charge de calcul des algorithmes mettant en oeuvre des procédures de mise en correspondance : approche du plus grossier au plus fin (*coarse-to-fine matching*). Une propriété essentielle d'une telle représentation est que les structures aux niveaux grossiers doivent être vues comme des simplifications de structures aux échelles plus fines - en d'autres termes, de nouvelles entités ne doivent pas apparaître.

2.5.1. Les pyramides

Outre les transformées orthogonales, les transformées en ondelettes ou pyramides QMF (*quadrature mirror filter*) qui permettent une décomposition du signal prenant en compte une notion d'orientation, une décomposition multi-

Dans ces conditions, la courbure gaussienne s'exprime par :

$$K = \frac{f_{uu}f_{vv} - f_{uv}^2}{(1 + f_u^2 + f_v^2)^2} \quad (\text{EQ 6})$$

Trois scénarios sont envisagés :

1. la direction verticale de l'engin est connue et l'échantillonnage des deux cartes est identique,
2. la direction verticale est inconnue mais l'échantillonnage est identique,
3. la direction verticale est inconnue et les deux cartes sont échantillonnées différemment, ce qui est le cas le plus réaliste et aurait pu être le cas le plus intéressant pour notre application. Toutefois, les auteurs doivent émettre quelques hypothèses avec notamment un re-échantillonnage des deux ensembles de données à la même période.

2.4 Une approche basée sur l'ensemble des données

L'algorithme proposé par Zhang [22] est basé sur la mise en correspondance itérative de deux ensembles de points 3D dans le contexte de navigation autonome d'un véhicule planétaire utilisant la vision. Il ne nécessite aucune connaissance *a priori* sur les images à recaler, mais suppose que le mouvement entre les deux vues successives est petit ou connu avec une assez bonne précision ; il utilise une technique qui s'appuie sur l'analyse statistique des distances d'appariements pour écarter les points erronés.

L'objectif est d'estimer de manière très précise le déplacement entre les deux vues afin de recaler deux cartes de profondeur.

L'algorithme peut se résumer ainsi :

on dispose en entrée de deux nuages de points 3D correspondant aux échantillons de deux surfaces et d'une estimée raisonnable initiale du déplacement entre les vues ; si $\{x_i\}$ est l'ensemble des points 3D de la première vue, on essaie d'apparier les x_i , après leur avoir appliqué l'estimée du mouvement (R, t) , avec leurs plus proches voisins de la seconde vue.

Itération jusqu'à convergence de l'estimation du mouvement :

- Pour chaque point de la première vue, trouver ses plus proches voisins dans la seconde par minimisation des distances,

L'algorithme se divise en trois étapes :

- calcul de la courbure Gaussienne à partir des données d'élévation,
- détermination des points de forte (ou faible) courbure par seuillage manuel,
- mise en correspondance de ces points.

Il s'agit, pour les auteurs, d'obtenir des points caractéristiques d'une surface qui soient invariants par isométrie (ce qui inclut l'invariance par rotation et translation). Si la courbure Gaussienne ne décrit pas complètement une surface, il reste que les points de forte courbure suffisent pour la reconnaître.

Le calcul de la courbure gaussienne :

Une carte de profondeur peut être décrite sous la forme :

$$X(u, v) = [u, v, f(u, v)]^T \quad (\text{EQ 3})$$

Dans ce cas, les dérivées partielles s'expriment par :

$$\left\{ \begin{array}{l} \frac{\partial X}{\partial u} = \left[1, 0, \frac{\partial f}{\partial u} \right]^T \\ \frac{\partial X}{\partial v} = \left[0, 1, \frac{\partial f}{\partial v} \right]^T \end{array} \right. \quad (\text{EQ 4})$$

$$\left\{ \begin{array}{l} \frac{\partial^2 X}{\partial u^2} = \left[0, 0, \frac{\partial^2 f}{\partial u^2} \right]^T \\ \frac{\partial^2 X}{\partial v^2} = \left[0, 0, \frac{\partial^2 f}{\partial v^2} \right]^T \\ \frac{\partial^2 X}{\partial uv} = \left[0, 0, \frac{\partial^2 f}{\partial uv} \right]^T \end{array} \right. \quad (\text{EQ 5})$$

en correspondance de ces modèles n'est pas adaptée au problème et au type de données.

Quant aux contours de même profondeur, ils apparaissent comme les primitives les plus intéressantes dans le cas de l'imagerie sous-marine ; en effet, les scènes observées présentent très peu de discontinuités, ce sont plutôt des surfaces lisses à partir desquelles il est facile d'extraire les lignes de même profondeur.

Le problème de recalage (*registration*) est de déterminer la matrice de la transformation $T(\theta, tx, ty)$ entre les deux systèmes de coordonnées définissant les deux images. Il s'agit là aussi d'une transformation dans un plan (trois paramètres) puisque l'information de profondeur est identique pour les deux systèmes et que l'on connaît tangage et roulis du bateau.

Dans ce cadre, les différentes étapes de la méthode proposée sont alors :

- re-échantillonnage régulier des données à partir d'une grille régulière,
- extraction des contours à partir de la grille régulière d'échantillonnage,
- codage des contours par un code de chaîne (différent de celui de Freeman dans le sens où les orientations des segments représentant le contour peuvent prendre un nombre infini de valeurs),
- détermination du recouvrement pour une paire de contours,
- calcul de la matrice de transformation pour chaque appariement de contour,
- test de cohérence pour chaque transformation ainsi calculée.

Les résultats présentés sont très convaincants. Toutefois, cette méthode est intéressante dans le cas où les images représentent des zones peu accidentées (alors les contours de même profondeur sont les primitives discriminantes).

2.3 Une approche basée sur des points spéciaux

Le problème est posé dans le cas de la navigation aérienne [21] : étant donné une carte d'élévation couvrant une zone de référence étendue et une carte d'élévation d'une petite partie de cette zone, il s'agit d'extraire les caractéristiques discriminantes et, en les mettant en correspondance, de localiser cette sous-région dans la carte de référence.

avec $r^2 = i^2 + j^2$, M le rayon du filtre et σ l'écart-type.

Désormais, on est ramené à un algorithme de mise en correspondance en dimension 2. L'étape suivante concerne l'extraction de points critiques le long de ces contours. Ils sont définis comme les extrema locaux de la courbure et supérieurs à seuil donné. Un contour est représenté par un codage de Freeman. A partir du code de Freeman, les points critiques sont extraits puis appariés : chaque point critique de la première carte (DEM) est comparé à l'ensemble des points critiques de la seconde (REM). On conclut à une hypothèse d'appariement lorsque l'erreur entre leurs voisins est petite.

Il est alors possible - à partir de chacun des appariements - de déterminer une transformation 2D (rotation et translation dans un même plan). Cette transformation hypothétique est appliquée à l'ensemble de la carte REM. La carte ainsi obtenue est «superposée» à la carte DEM ; si l'alignement des deux cartes est supérieur à 30%, la transformation est retenue. Le critère d'alignement est le nombre de points critiques de la carte REM transformée qui se trouvent près (5 pixels) de ceux de la carte DEM avec la même orientation (à 90 degrés près). Les résultats obtenus sur des données réelles montrent, selon les auteurs, que l'utilisation des contours de falaise est une technique efficace et robuste pour la mise en correspondance.

2.2.2. L'approche de Kamgar-Parsi, Jones et Rosenfeld

Cette approche [20] est sensiblement identique à la précédente : il s'agit de mettre en correspondance des cartes de distances sous-marines acquises par un sondeur multi-faisceaux (bathymétrie) pour lesquelles les caractéristiques discriminantes sont rares. Pour cela, les auteurs montrent que les primitives les plus appropriées à ce problème sont les contours de même profondeur.

A cet endroit, il est intéressant de noter les raisons qui ont conduit à ce choix. Les *templates* ou sous-images engendrent des temps de calcul considérables dus au fait que toutes les orientations et positions possibles d'un *template* provenant de la première image doivent être testées dans la seconde image afin de déterminer l'appariement correct ; de plus, la résolution est souvent mauvaise sur des images bathymétriques ; l'extraction de contours pour lesquels les niveaux de gris subissent de brusques variations (*edges*), technique courante pour les systèmes de stéréovision pour lesquels l'orientation relative des deux vues est connue, n'est pas appropriée ici et conduit à un grand nombre d'appariements erronés ; pour des raisons analogues (absence de classes ou de primitives singulières et discriminantes, segmentation complexe), la description des images par des fonctions analytiques puis la mise

Toutefois la corrélation reste envisageable combinée à une approche hiérarchique. Dans [16], le problème consiste à conduire automatiquement un senseur TV depuis une plate-forme aérienne vers un objet connu, fixe et déterministe. L'algorithme de mise en correspondance entre le modèle de l'objet et l'image doit s'affranchir des différences de facteur d'échelle entre le masque de l'objet (image binaire des contours de l'objet) et l'image binaire des contours, et des différences de rotation. Dans ces conditions, la mesure de similitude choisie est la fonction d'intercorrélation sur deux images binaires. Pour la rendre indépendante des problèmes de changement d'échelles et de rotations, plusieurs masques de l'objet sont stockés. Cette approche est bien entendue limitée : il s'agit là d'un problème de reconnaissance de formes *binaires*.

2.2 Les approches basées sur les contours

2.2.1. L'approche de Rodriguez et Aggarwal

Cette approche [17][18][19] est basée sur la mise en correspondance de contours. Il s'agit d'un algorithme de recalage d'une séquence d'images aériennes avec une carte 3D de référence embarquée (*Digital Elevation Map*), l'objectif étant l'élaboration d'un système passif de navigation aérienne, c'est-à-dire la détermination de la trajectoire de l'engin (position et cap).

En entrée, on dispose d'une séquence d'images optiques monoculaires et d'une carte d'élévation de référence. Préalablement, il faut que ces deux types de données soient « convertis » dans une représentation identique. L'idée est de transformer les images optiques en modèles numériques de terrain (MNT) par des procédés de stéréovision (*Recovered Elevation Map*) : chaque paire d'images consécutives dans la séquence aérienne est traitée comme une paire d'images stéréo. En outre, pour réduire la taille des données et pour faciliter l'algorithme de recalage, les deux cartes de profondeur (DEM et REM) sont transformées en cartes de falaises (*cliffs maps*). Le problème est alors la mise en correspondance par appariement de points critiques le long des contours.

Les contours sont obtenus par convolution des MNT avec un filtre LoG (*Laplacian-of-Gaussian*) puis par détection des passages par zéro.

Dans le cas discret, le filtre LoG est de la forme :

$$\nabla^2 G(i, j) = \frac{1}{2\pi\sigma^2}(r^2 - 2\sigma^2)\exp\left(-\frac{r^2}{2\sigma^2}\right) \quad (\text{EQ 2})$$

pour tout $i \in [-M, M]$ et $j \in [-M, M]$,

Chapitre 1 : Un premier état de l'art.

2 La mise en correspondance

Notre problème consiste en la mise en correspondance de cartes bathymétriques sous-marines. Ces cartes peuvent être vues comme des images de profondeur. Les approches que nous exposons dans le chapitre suivant nous paraissent intéressantes notamment du point de vue du type de données qui sont traitées, compte tenu de la qualité de celles dont nous disposons, l'imagerie SONAR étant ce qu'elle est ! Notons que la littérature est relativement pauvre quant au traitement de ces images ...

2.1 Une approche générale : la corrélation.

La corrélation est l'approche statistique générale de mise en correspondance. En elle-même, ce n'est pas une méthode de mise en correspondance, mais une mesure de similarité entre une image et une zone d'intérêt (ou *template*). Soit T la sous-image que l'on souhaite retrouver dans l'image I .

Le score de corrélation croisée normalisée est, pour toute translation (u, v) :

$$C(u, v) = \frac{\sum_x \sum_y T(x, y) I(x + u, y + v)}{\sqrt{\left[\sum_x \sum_y I^2(x + u, y + v) \right]} \sqrt{\left[\sum_x \sum_y T^2(x, y) \right]}} \quad (\text{EQ 1})$$

Le score maximal $C(u_0, v_0)$ renvoie à la translation (u_0, v_0) de T par rapport à I . Outre la cas des translations, ce score peut être généralisé aux cas pour lesquels T a subi une rotation et un changement d'échelle ; il sera alors calculé, non seulement pour toutes les translations, mais aussi toutes les rotations et toutes les échelles envisageables. Il est évident que l'algorithme reste coûteux en temps de traitement. Généralement, une telle méthode s'applique lorsque la transformation entre T et I , de type similitude ou affine, est petite. D'autre part, si cette approche ne nécessite pas *a priori* de pré-traitement des images, elle demeure sensible au bruit et aux distorsions, aux déplacements non-linéaires, aux occlusions...

On lui préfère dans certains cas la Transformation de Hough [15] : elle permet alors de s'affranchir de problèmes tels que le changement d'orientation des objets dans le temps, ou du manque d'informations partiel (occlusion) ; d'autre part, la transformation de Hough se généralise au cas des séquences d'images pour l'estimation de mouvement.

permis de valider expérimentalement la mise en correspondance de cartes de profondeur pour déterminer la position absolue du véhicule.

Le scénario que nous envisageons concerne la mise en correspondance de cartes bathymétriques prises localement par un engin autonome ou remorqué avec une carte globale embarquée réalisée préalablement par un navire de surface à une attitude supérieure, avec une couverture plus importante.

Une telle étude peut alors se diviser en trois parties essentielles :

- Analyse d'images sonar et de profils bathymétriques. Un des critères essentiels pour assurer l'embarquabilité d'une solution opérationnelle d'un système de navigation référencé terrain réside dans la taille des données à traiter et la complexité des algorithmes associés.
- Classification et analyse multi-échelles. Cette analyse pourra être vue comme un traitement préalable des données.
- Mise en correspondance et estimation. La mise en correspondance n'est pas à étudier uniquement suivant des critères spatiaux mais il faut tenir compte des aspects temporels (mouvement et séquences d'images).

Nous présentons ici un premier état de l'art concernant ces trois points. Dans une seconde partie, nous exposerons deux applications, la première visant à valider l'algorithme proposé par Rodriguez dans [19] sur nos propres données bathymétriques, la seconde proposant une méthode d'estimation de mouvement tridimensionnel par mise en oeuvre d'un filtre de Kalman, celui-ci prenant part intégrante dans l'algorithme de Zhang [23] quant au recalage de nuages de points 3D.

1 Introduction

La navigation est définie comme l'action de conduire d'un point à un autre un véhicule et d'en déterminer la position à chaque instant. Matière essentielle pour les robots mobiles autonomes.

Elle a été étudiée en détail dans les domaines maritimes, spatiaux ou aériens. Aussi est-il surprenant de voir combien cette question reste problématique mais néanmoins stratégique dans le monde de la robotique autonome, étant donné l'ensemble des connaissances et l'expérience dans les communautés maritimes et aérospatiales.

C'est que la recherche dans le domaine des véhicules autonomes tend à éliminer peu à peu l'utilisation de balises artificielles - qu'elles soient actives ou passives - placées dans l'environnement du robot à des fins de localisation et de navigation. La raison essentielle réside dans l'essor de l'intelligence artificielle, dans le but de développer des engins mobiles «intelligents». Dans cette optique, un système de navigation pour véhicule autonome se doit d'identifier, de reconnaître et d'utiliser les différents amers et autres repères de son environnement.

La navigation autonome des véhicules sous-marins passe par une estime optimale et des moyens de recalage asynchrones, usuellement fournis par des systèmes de navigation acoustique de type Base Longue.

Une classification habituelle distingue localisation relative ou à l'estime et localisation absolue, suivant la nature des données perçues par les capteurs :

- lorsqu'il s'agit de capteurs de vitesse, d'accélération ou d'attitude (compas, loch, profondimètre, etc.), lorsque ceux-ci réagissent aux mouvements relatifs du véhicule et qu'ils ne permettent pas de donner directement un modèle de l'environnement (capteurs non cognitifs), on parlera de *localisation relative* ou à l'estime ; le positionnement se fait alors «pas à pas» sans référence à un modèle de l'environnement,
- lorsqu'ils permettent de percevoir directement et individuellement l'environnement et de le modéliser (capteurs cognitifs : vidéo, sonar, etc.), on parle alors de *localisation absolue* ou avec modèle de référence ; le positionnement résulte de la mise en correspondance du modèle fourni par les capteurs et du modèle global de l'environnement connu *a priori*.

En navigation locale, on utilise le sonar pour la recherche et la poursuite de cibles fixes ou mobiles ; en navigation long terme, des travaux récents ont

LISTE DES FIGURES

<i>Figure 1</i> - Estimation par approche hiérarchique	23
<i>Figure 2</i> - Le Sondeur EM12 multi-faisceaux du navire l'Atalante	35
<i>Figure 3</i> - Méthode d'acquisition	37
<i>Figure 4</i> - Intersection entre émission et réception	38
<i>Figure 5</i> - Le canyon du Var (EM12) : Visualisation du relief (a) et isobathes (b)	39
<i>Figure 6</i> - Niveaux de gris de la carte bathymétrique du canyon du Var .	40
<i>Figure 7</i> - Carte de contours	41
<i>Figure 8</i> - Les points critiques	43
<i>Figure 9</i> - Le système de navigation du robot autonome Blanche	44
<i>Figure 10</i> - Synoptique du système de navigation référencée terrain	44
<i>Figure 11</i> - Filtre de Kalman : Bloc Diagramme	50
<i>Figure 12</i> - Les scènes de rocher	59
<i>Figure 13</i> - Comparaison de la convergence des 2 approches (test 2) avec Del- ta-X = $\ X_i - X_{i-1}\ / \ X_i\ $	61
<i>Figure 14</i> - Comparaison de la convergence des 2 approches (test 3) . . .	62

Chapitre 2 : Quelques résultats appliqués à nos données.

5	Les caractéristiques des sondeurs multi-faisceaux EM12 et EM1000	35
5.1	Représentation et traitement des données	39
5.2	Perspectives	43

Chapitre 3 : Estimation de mouvement 3D par filtrage de Kalman

6	Les algorithmes de mise en correspondance et d'estimation de mouvement	45
7	Le filtrage de Kalman	46
7.1	Le filtre de Kalman et l'estimation de mouvement	46
7.2	Les équations générales du filtre de Kalman et du filtre de Kalman étendu	47
7.2.1	Le filtre de Kalman	48
7.2.2	Le Filtre de Kalman Etendu	51
8	La représentation du mouvement.	53
8.1	Le mouvement rigide	53
8.2	La représentation choisie	53
9	L'estimation du mouvement 3D	54
9.1	Calcul des dérivées	55
9.2	Le traitement des incertitudes	55
9.3	Condition de convergence	56
10	Les résultats	56
10.1	sur des données synthétiques	57
10.2	sur des données réelles	58

Bibliographie

TABLE DES MATIERES

1	Introduction	9
	<i>Chapitre 1 : Un premier état de l'art.</i>	
2	La mise en correspondance	11
2.1	Une approche générale : la corrélation.	11
2.2	Les approches basées sur les contours	12
2.2.1	L'approche de Rodriguez et Aggarwal	12
2.2.2	L'approche de Kamgar-Parsi, Jones et Rosenfeld	13
2.3	Une approche basée sur des points spéciaux	14
2.4	Une approche basée sur l'ensemble des données	16
2.5	Vers une approche multi-échelles	17
2.5.1	Les pyramides	17
2.5.2	La représentation espace-échelle (scale-space)	20
2.6	Mise en correspondance et multi-résolution	22
3	Analyse de séquences d'images	23
3.1	Mise en correspondance et estimation de mouvement	24
3.2	Estimation du mouvement d'un objet 3D se déplaçant dans un environnement 3D	25
3.3	La notion de flot optique	27
3.4	L'estimation robuste	28
3.5	Une approche d'estimation robuste multi-échelles de mouvement	30
4	Vers la navigation référencée terrain	32
4.1	La vision pour le contrôle des véhicules sous-marins	32
4.2	L'approche de Bergem.	33
4.3	L'approche de Cristi	34

Toward terrain based underwater navigation by bathymetric profiles matching

Abstract: Navigation is usually taken to mean the conduct of a vehicle from one point to another, and, in order to do this, the determination of its position, course and velocity at any moment. The techniques may be divided in two main categories: relative navigation and absolute navigation, depending on whether they determine the position by dead-reckoning, from velocity or acceleration measurements or whether they measure an absolute position with respect to artificial or natural beacons at known position. Our problem concerns sonar based navigation for absolute location of autonomous underwater vehicles. It deals with matching a local depth maps against a large, on-board, low resolution reference map. Including information from all available sensors (inertial sensors, velocity, etc.) position error may significantly be reduced. Such a problem is divided in three significant parts: bathymetry profile analysis, classification and multi-scale analysis, registration and estimation. Our contribution will focus on matching and estimation. We propose here a survey of the main works reported in the literature. Then we present two applications, a first approach of feature extraction using real bathymetric data and estimation of 3D-motion using Kalman filtering.

Key-words: Navigation, Sonar, Matching, Multi-Scale Analysis, Motion Estimation.

Vers la navigation référencée terrain par mise en correspondance de cartes bathymétriques

Laurence Lucido *

Thème 3

Intéraction homme-machine, images, données, connaissances

Projet RobotVis

Rapport de recherche n°3211- Juillet 1997

Résumé : La navigation est définie comme l'action de conduire d'un point à un autre un véhicule, et d'en déterminer la position à chaque instant. Une classification habituelle consiste à distinguer la localisation relative ou «à l'estime» de la position du sous-marin, à l'aide des capteurs de vitesse, de cap, d'immersion, et la localisation absolue par rapport à des balises actives ou à des amers passifs. Le scénario que nous envisageons concerne essentiellement ce dernier point, mais il s'agira d'optimiser le système en utilisant notamment l'ensemble des informations provenant des capteurs d'estime. Notre problème concerne l'utilisation du sonar bathymétrique pour la navigation et le pilotage autonome des engins sous-marins : il s'agit de mettre en correspondance des cartes bathymétriques prises localement par un engin autonome ou remorqué avec une carte globale embarquée réalisée préalablement par un navire de surface à une attitude supérieure, avec une couverture plus importante. Une telle étude peut se diviser en trois parties essentielles : analyse de profils bathymétriques, classification et analyse multi-échelles, mise en correspondance et estimation. Notre contribution portera principalement sur ces derniers aspects. A ce sujet, nous présentons dans ce rapport un premier état de l'art quant aux points clés de cette étude. Nous exposons deux applications, l'une concernant le traitement d'une carte bathymétrique réelle, l'autre s'intéressant à l'estimation de mouvement par filtrage de Kalman étendu.

Mots-clé : Navigation, Sonar, Mise en Correspondance, Analyse Multi-Echelle, Estimation de Mouvement.

(Abstract: pto)

* Email : lucido@ifremer.fr

IFREMER - Laboratoire de Robotique Sous-Marine -
ZP de Brégaillon - BP 330 - 83507 La Seyne-Sur-Mer (France)
Téléphone : 04 94 30 44 18 - Télécopie : 04 94 87 83 07



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Vers la navigation référencée terrain
par mise en correspondance de cartes
bathymétriques

Laurence LUCIDO

N° 3211

Juillet 1997

Thème 3

Intéraction homme-machine, images,
données, connaissances

Projet RobotVis

*R*apport
de recherche

1997