

Du 2D au 3D : Imagerie stéréoscopique et estimation du relief(*)

Julien Sénégas(**)

Le mois de mai 2002 a vu la mise sur orbite du satellite SPOT 5 (Satellite Pour l'Observation de la Terre) avec pour objectif « de constituer en cinq ans un modèle numérique de terrain mondial de 30 à 50 millions de km² correspondant au tiers des terres émergées » (Jean-Marc Nasr, PDG de SPOT Image, Libération, 3 mai 2002). Un Modèle Numérique de Terrain (MNT) est une carte où la variable exprimée est l'altitude au point considéré. Si un point de la surface est représenté par ses coordonnées (x, y, z) , il s'agit donc dans une description mathématique de l'application $(x, y) \rightarrow z(x, y)$. Les utilisateurs de ces bases de données cartographiques sont multiples : l'armée est un des acteurs historiques de ce marché, mais il faut également citer l'aviation civile, les collectivités locales qui désirent gérer l'aménagement de leur territoire, et un certain nombre d'activités économiques de haute technologie comme les télécommunications qui ont besoin de données récentes pour optimiser la position de leurs relais de téléphonie mobile, les systèmes de navigation embarqués, ...

L'imagerie stéréoscopique constitue avec le radar une des principales technologies permettant de recueillir des données utiles à la reconstitution 3D et de calculer automatiquement des MNTs. Le principe général, utilisé en photogrammétrie et en vision par ordinateur, est relativement simple et s'appuie sur une analogie avec le système visuel humain : à partir de deux vues « planes » d'une scène⁽¹⁾, il est possible de reconstituer sous certaines conditions une information tri-dimensionnelle (Figure 1). Dès lors, en utilisant des images satellitaires ou aériennes recouvrant de grandes zones géographiques (120 km par 60 km avec une résolution de 5 m pour SPOT 5), des modèles numériques de terrain peuvent être produits de manière automatique et en un temps relativement court. Dans des pays comme la France ou les États-Unis, l'intérêt des données issues de l'imagerie satellitaire ou aérienne est principalement d'actualiser les bases de données géographiques déjà existantes, que l'on s'intéresse par exemple à l'évolution des milieux urbains ou à la gestion de l'environnement ; mais en Asie ou en Amérique du Sud, bien souvent, aucune donnée n'est disponible et ces nouvelles techniques sont donc essentielles pour constituer en un temps limité un catalogue important.

(*) Travail réalisé en collaboration avec le Centre de Géostatistique de l'École des Mines de Paris, 35 Rue St Honoré, 77300 Fontainebleau, France, et Istar, 2600 Route des Crêtes, BP 282, 06905 Sophia Antipolis, France.

(**) Adresse actuelle : Philips Forschungslaboratorien, Roentgenstrasse 24, 22335 Hambourg, Allemagne. E-mail : julien.senegas@philips.com

(1) Pour des exemples de couples stéréoscopiques satellitaires, nous renvoyons le lecteur à l'adresse Internet : www.spotimage.fr

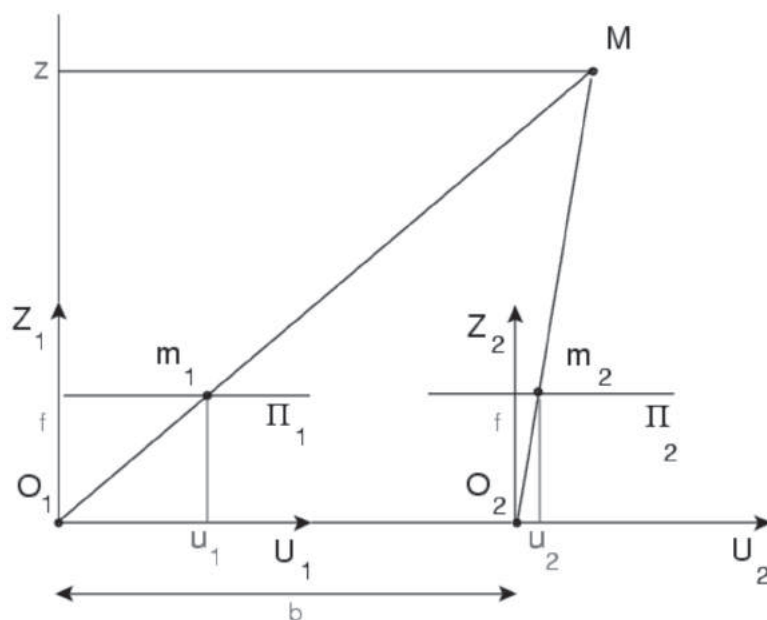


Figure 1 : Représentation schématique d'un système stéréoscopique dans le cas de caméras parallèles. Chaque caméra est symbolisée par son centre optique O et son plan image P . L'ordonnée z du point M peut être calculée à partir des abscisses u_1 et u_2 des deux points homologues m_1 et m_2 , de la distance focale f et de la distance b entre les deux centres optiques.

Problématique générale

ISTAR (Imagerie STéréoscopique Appliquée au Relief), société française créée en 1988 par des chercheurs issus de l'INRIA, est un des principaux acteurs mondiaux sur le marché des bases de données géographiques numériques. Elle utilise actuellement principalement deux types de données : des images obtenues par les capteurs embarqués sur les satellites SPOT ou LANDSAT et des images aériennes haute résolution produites par une caméra spécialement développée par le centre d'études spatiales allemand, le DLR. À partir de ces images sont produits de manière quasiment automatique des MNTs, à différentes résolutions selon le capteur utilisé, et des ortho-images (images corrigées des déformations induites par le relief).

Le travail présenté ici répond à une problématique soumise par ISTAR : *comment chiffrer l'incertitude associée à un modèle numérique de terrain calculé à partir d'un couple stéréoscopique donné ?* Cette question est en effet primordiale dès que le MNT est utilisé à des fins industrielles ou stratégiques : il faut pouvoir assurer la fiabilité d'un système ou la validité d'une décision en tenant compte des erreurs possibles. Notons que ce problème nécessite également de définir les outils appropriés pour décrire l'incertitude. Il s'agit par exemple de calculer un intervalle de confiance de l'altitude du terrain à un niveau de risque fixé.

La difficulté principale de l'étude réside dans l'absence d'information autre que le couple stéréoscopique ; en particulier, on renonce à recourir à des mesures directes

de l'altitude, car dans de nombreux cas, celles-ci ne sont pas disponibles. À partir d'un même couple stéréoscopique, dans notre cas d'images SPOT, le problème est donc de calculer le MNT correspondant et d'en estimer la précision.

Le problème de la vision stéréoscopique a suscité un grand nombre de travaux, qui ont abouti au développement d'un nombre tout aussi grand d'algorithmes. De manière générale, la résolution du problème est scindée en trois étapes distinctes [2] : la calibration, qui consiste à calculer les paramètres géométriques du système stéréoscopique, la mise en correspondance, dont l'objet est d'identifier les couples de points images provenant d'un même point physique, puis la triangulation spatiale, à savoir le calcul en lui-même des coordonnées spatiales des points de la scène observée. Calibration et triangulation reposent essentiellement sur des relations géométriques dépendant de la nature du capteur utilisé, et sont actuellement bien maîtrisées. En revanche, la mise en correspondance est toujours un domaine actif de recherche [1]. C'est à cette étape que nous allons nous intéresser par la suite. De fait, actuellement, aucun des algorithmes de calcul n'est en mesure de donner une estimation précise de la fiabilité de la mise en correspondance. En réalité, ce problème est rarement abordé en tant que tel : on se contente généralement de chiffrer la précision d'un algorithme à partir d'un jeu de données de référence.

Le point de vue bayésien

Revenons au problème de la mise en correspondance et considérons deux acquisitions d'une même scène. Un point physique, en l'absence d'occlusions, est projeté dans chacune des deux images : les deux projections forment un couple de points homologues. La mise en correspondance consiste alors à déterminer l'ensemble des couples de points homologues à partir des deux images recueillies. Il s'agit donc d'un problème inverse dont la résolution dépend fortement de la qualité de l'information disponible, à savoir la radiométrie des images du couple. Les erreurs d'appariement ne dépendent pas du système d'observation, mais du contenu du couple et donc de la nature de la scène observée ; elles ne sont pas systématiques. Elles ont en revanche un caractère aléatoire : non pas que le phénomène étudié en lui-même soit aléatoire (cela n'aurait que peu de sens d'avancer que l'élévation d'un terrain est le résultat d'une expérience aléatoire), mais parce que l'information utilisée est bruitée et par nature ambiguë (Figure 2). Il faut bien comprendre que dans le contexte de la reconstruction 3D, il existe une incertitude intrinsèque au couple stéréoscopique, une partie de l'information n'étant pas accessible. C'est cet aspect que nous proposons de traiter.

Adoptons un point de vue bayésien et reformulons le problème. On dispose d'une information de nature radiométrique, le couple stéréoscopique, notée Y par la suite. À partir d'une réalisation particulière y de ce couple, on désire reconstituer la géométrie Z de la scène observée. La connaissance que l'on peut avoir de Z à partir de Y est entièrement décrite par la probabilité *a posteriori* de Z sachant $Y = y$. En d'autres termes, parmi toutes les configurations Z possibles, on s'intéresse à celles qui se réalisent lorsque y est observé. L'incertitude quant à l'estimation de Z à partir de Y est contenue dans cette probabilité *a posteriori*.

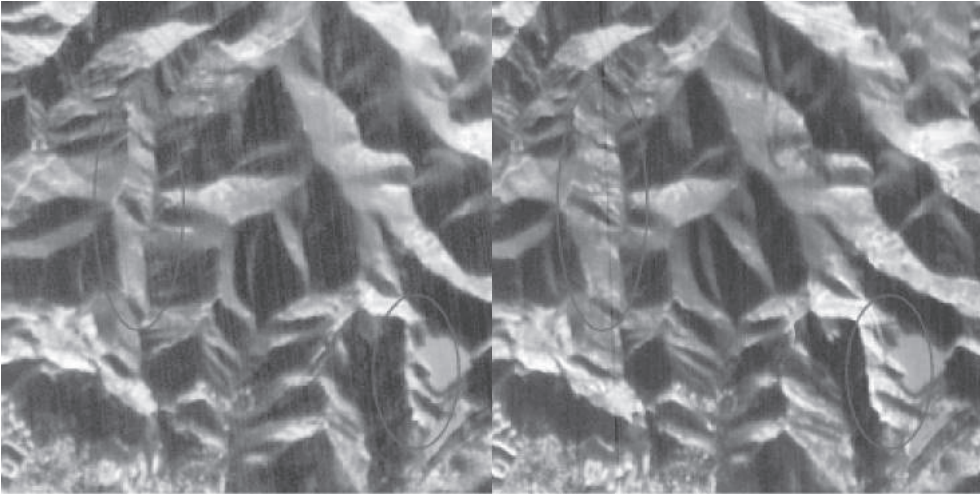


Figure 2 : Exemple de déformations géométriques importantes dans un couple stéréoscopique de région montagneuse. À noter également la présence de zones avec un contraste faible (ombres dues au relief). Dans ces deux cas, l'information radiométrique ne permet pas d'établir de manière univoque la correspondance. Des hypothèses supplémentaires (information *a priori*) sont requises pour calculer une solution.

Soit un espace probabilisé E , dans lequel la variable aléatoire Z prend ses valeurs. Notons $\pi_Z(\cdot|y)$ la loi *a posteriori* de Z sachant $Y = y$. Si nous notons \hat{z} une estimation de la géométrie, calculée à partir d'un algorithme quelconque, il est possible de calculer par exemple la probabilité de dépassement d'un certain seuil s :

$$P(|Z - \hat{z}| \geq s) = E_{\pi_Z(\cdot|y)}(1_{|Z - \hat{z}| \geq s}) = \int 1_{|Z - \hat{z}| \geq s} \pi_Z(z|y) dz.$$

Inversement, si l'on se fixe un risque α , le problème peut être de déterminer l'intervalle de confiance $[z_{\text{inf}}; z_{\text{sup}}]$ associé, i.e. tel que :

$$P(Z \in [z_{\text{inf}}; z_{\text{sup}}]) \geq 1 - \alpha.$$

Nous venons donc de donner une formulation mathématique au problème initial. Le cadre est un cadre probabiliste, bayésien au sens où nous nous intéressons à la loi *a posteriori* $\pi_Z(\cdot|y)$ de la scène Z connaissant le couple stéréoscopique $Y = y$. On pourrait également utiliser le terme de loi conditionnelle. Ce cadre probabiliste nous permet de quantifier de manière simple l'incertitude relative à Z .

Vers les méthodes de Monte-Carlo

Si la formulation du problème repose sur le choix de la distribution $\pi_Z(\cdot|y)$, sa résolution implique le calcul d'intégrales multiples. Or, la dimension du problème est telle (la taille d'une scène SPOT est de l'ordre de 36×10^6 pixels) qu'un calcul direct ne peut être mis en œuvre. Une solution bien connue est celle apportée par les méthodes dites de Monte-Carlo, largement utilisées autant en analyse numérique qu'en calcul de probabilité. Elles consistent à générer des réalisations de la loi $\pi_Z(\cdot|y)$

et à remplacer le calcul de l'intégrale par une moyenne arithmétique. Pour une loi π et une fonction g π -mesurable, on a en effet :

$$E_{\pi}(g) = \int g(z)\pi(z)dz \approx \frac{1}{n} \sum_{i=1}^n g(z_i)$$

où $z_i, i = 1, \dots, n$ sont des réalisations de la loi π .

Encore faut-il pouvoir simuler la loi désirée... Deux approches existent. Pour certaines lois, il existe un algorithme de simulation directe, par exemple par inversion de fonction de répartition ou par acceptation/rejet. C'est le cas notamment des modèles gaussiens, pour lesquels un grand nombre d'algorithmes fréquemment utilisés en géostatistique existent. En revanche, lorsque la loi ne peut se déduire d'une loi aisément simulable, ce qui est souvent le cas lorsque l'on considère des lois conditionnelles, il semble que seules des méthodes itératives permettent d'apporter une réponse. Il s'agit principalement de générer une chaîne de Markov admettant pour loi stationnaire la loi à simuler [4]. Parmi les algorithmes les plus utilisés, citons en particulier l'échantillonneur de Gibbs [3] (Figure 3) et l'algorithme de Metropolis-Hastings [5].

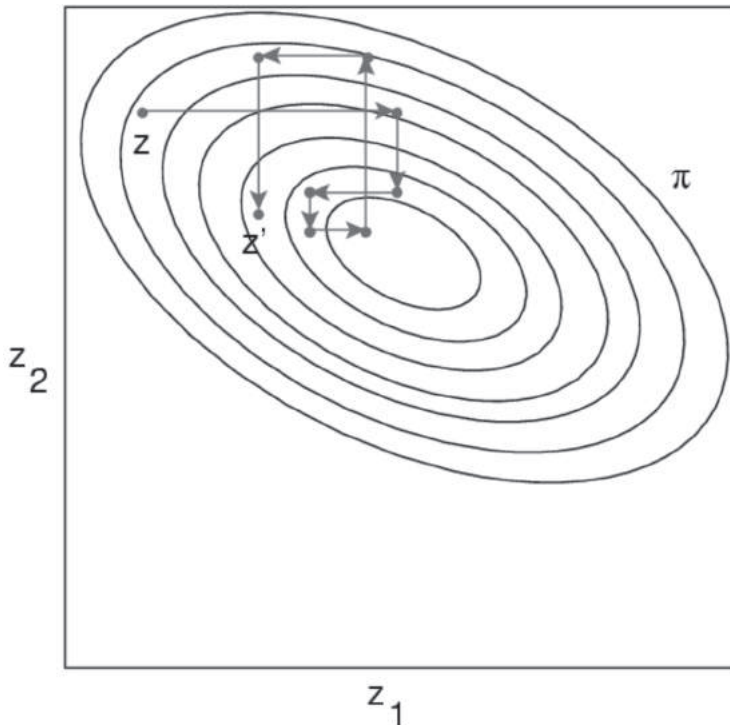


Figure 3 : Exemple de transitions générées par l'échantillonneur de Gibbs pour une loi bivariable π , représentée en lignes de niveau. L'état z' est obtenu après quatre cycles de l'échantillonneur. Lors d'un cycle, chaque composante z_i est relaxée à l'aide d'une simulation selon la loi conditionnelle $\pi(z_i | z_j, j \neq i)$.

En statistique, le problème de l'étude de lois de probabilité par simulations markoviennes et intégrales de Monte-Carlo, couramment désigné par l'expression Markov chain Monte-Carlo (MCMC), est un des domaines de recherche les plus actifs, surtout depuis le début des années 1990 où ces méthodes ont fait leur apparition en analyse bayésienne. Ce qui apparaissait initialement comme un problème propre au domaine de la vision par ordinateur fait donc en réalité référence à une problématique beaucoup plus générale et nécessitant un important développement probabiliste.

Sur le choix du modèle

Nous sommes passés un peu vite sur la formulation du problème pour déjà nous intéresser à sa résolution. Or, une grande partie du travail porte sur le choix du modèle probabiliste $\pi_z(\cdot|y)$ susceptible de décrire le paradigme de la vision stéréoscopique. Si l'approche proposée ici et l'utilisation des méthodes MCMC sont tout à fait générales, le modèle choisi est forcément spécifique et tourné vers l'application considérée. Notamment, le choix de travailler avec des images SPOT a une incidence sur la nature des scènes observées : à cette résolution, seules les constructions urbaines importantes et les variations sensibles du relief peuvent être distinguées.

Il est primordial de comprendre que l'utilisation d'un modèle est forcément subjective. Il est notamment toujours beaucoup plus aisé d'en réfuter que d'en mettre en évidence la validité. C'est pourquoi nous utilisons dans la suite l'expression *choix* du modèle, afin de signaler qu'il s'agit là d'un parti pris non négligeable. Cette subjectivité du modèle ne rend d'ailleurs que plus nécessaire les étapes de validation des résultats. Enfin, pour paraphraser l'adage matheronien *il n'y a pas de probabilité en soi, il n'y a que des modèles probabilistes* [6], gardons à l'esprit que les calculs de probabilité qui sont l'objet de ce travail ne sont valables que dans le cadre du modèle choisi, et n'ont aucune signification propre.

Apports

Les différents aspects du problème que nous venons de présenter sont traités et présentés dans [7] et [8]. Deux apports originaux peuvent être dégagés. Le premier est plutôt d'ordre conceptuel : il s'agit de la formulation mathématique du problème posé en termes d'échantillonnage d'une loi de probabilité. Cette approche, qui apparaît naturelle dès que l'on adopte une description bayésienne du problème, est cependant relativement nouvelle en vision par ordinateur et exige un changement de point de vue. Il y a en effet un pas important à franchir dans l'appréhension du réel entre une vision déterministe, où la solution est forcément unique, et une vision probabiliste, où c'est au contraire un ensemble de solutions que l'on envisage. Même si les méthodes bayésiennes sont employées couramment en vision par ordinateur, elles restent avant tout un outil de calcul servant à estimer la géométrie (ou toute autre propriété) de la scène. Autrement dit, bien qu'on utilise la loi de probabilité pour formuler le problème, bien souvent n'est retenue qu'une caractéristique de celle-ci (sa moyenne ou son maximum). La multiplicité des solutions continue de

gèner... Or c'est justement cette multiplicité qui est au cœur du problème qui nous intéresse.

Le second apport est plus théorique et concerne le développement d'algorithmes de simulation par chaînes de Markov. En particulier, nous proposons une extension de l'algorithme de Metropolis-Hastings en adoptant un point de vue multi-dimensionnel et développons plusieurs algorithmes dans un cadre gaussien. Certains aspects font également appel aux méthodes d'échantillonnage d'importance pour diminuer les temps de calcul. Ces algorithmes ont un intérêt tout à fait général dans le domaine des statistiques spatiales.

Bibliographie

- [1] M.Z. Brown, D. Burschka et G.D. Hager. *Advances in Computational Stereo. IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25 (8), 2003.
- [2] O. Faugeras. *Three-Dimensional Computer Vision*. Massachussets Institute of Technology, 1993.
- [3] S. Geman et D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE Transitions on Pattern Analysis and Machine Intelligence*, 6 (6), 1984.
- [4] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Chapman and Hall, 1996.
- [5] W.K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57 : 97-109, 1970.
- [6] G. Matheron. *Estimating and choosing – An Essay on Probability in Practice*. Springer-Verlag, Berlin, 1989.
- [7] J. Sénégas. A Markov chain Monte Carlo Approach to stereovision. In A. Heyden, G. Sparr, M. Nielsen, et P. Johansen, éditeurs, *European conference on Computer Vision*, volume 3, pages 97-111. Springer-Verlag, Berlin-Heidelberg, 2002.
- [8] J. Sénégas. *Méthodes de Monte Carlo en vision stéréoscopique – Application à l'étude de modèles numériques de terrain*. Thèse, École des Mines de Paris, 2002.