

B. Quelques actions marquantes

B.1 Appariement de clichés terrestres guidé par un modèle : application à la construction de mosaïque de linéaire de rue

Jocelyn Marchadier

Introduction

La motivation initiale de notre étude est le recalage d'un modèle tridimensionnel de type TRAPU et de clichés terrestres haute résolution. Les paramètres de prises de vue ainsi obtenus permettraient l'obtention d'un modèle tridimensionnel photoréaliste. Cette application nécessite la résolution de plusieurs problèmes, dont la mise en place et la fusion de textures, nécessaires à la gestion des occlusions. Certains de ces problèmes liés à la précision géométrique du modèle TRAPU ont été présentés dans le précédent Bulletin d'Information de l'IGN : n° 71 (1999/3).

Nous pouvons mettre en évidence la différence de géométrie entre le modèle initial, reconstruit à partir de photographies aériennes et la réalité présente sur les images. Ce problème, gênant lors du rendu d'un modèle texturé, est crucial pour l'obtention d'un algorithme d'appariement automatique robuste. Seule une compensation globale permettrait d'accroître la robustesse et la fiabilité du procédé.

La chaîne de traitement automatique permettant la mise en place des clichés peut être décomposée en plusieurs étapes :

- mise en correspondance de ces primitives entre images consécutives ;
- mise en place robuste de la bande d'images ;
- mise en correspondance de primitives image avec le modèle ;
- calcul d'une compensation globale robuste.

Nous présentons ici un ensemble de méthodes spécifique au contexte de notre étude qui peuvent être mise en œuvre pour résoudre les trois premières étapes décrites ci-dessus. Ces méthodes sont fiables et efficaces, ne nécessitant que peu de temps de calcul. À l'issue de la chaîne de traitement proposée, une mosaïque permettant l'étude d'un linéaire de rue peut être présentée, formant ainsi un produit dérivé de la chaîne de traitements.

La première partie de cet article montre l'ensemble des données que nous avons utilisées. La deuxième partie est consacrée à un algorithme d'extraction de primitives que nous avons publié et une méthode fiable d'extraction de points de fuite.

Une méthode efficace d'obtention d'appariements initiaux est proposée dans la troisième partie de cet article. La quatrième partie met en évidence l'utilité de la connaissance des points de fuite principaux des clichés terrestres pour filtrer les appariements initiaux. La cinquième partie de l'article présente une petite mosaïque conçue de manière rudimentaire qui démontre les fonctionnalités des traitements.

Données initiales

Le site choisi pour nos tests est la rue du faubourg Saint Antoine à Paris. Le modèle TRAPU utilisé est d'une précision de 20 cm. Seules les gouttières ont été mesurées, le volume des bâtiments ayant été reconstruit par extrusion du périmètre des toits au niveau du sol, qui est ici modélisé par un plan.

Un ensemble de clichés des façades des bâtiments a été réalisé à l'aide d'un appareil photographique 24 x 36, doté d'un objectif grand angle d'une focale de 20 mm. Ils forment une bande de 80 % de recouvrement, la base est d'approximativement six mètres. Ils ont été numérisés avec un scanner Helava d'une bonne qualité radiométrique, puis sous-échantillonnés. La taille du pixel des images utilisées est de 40 µm, soit cinq centimètres sur une façade.

Les clichés sont fortement déformés et offrent la particularité d'avoir deux points de fuite principaux.

Les paramètres internes des prises de vues ont été calculés en utilisant la même méthode que celle qui a été utilisée pour le chantier d'Angkhor Vat.

À l'aide des données initiales des prises de vue, et de la position de la première prise de vue, il est possible d'obtenir une approximation des paramètres des prises de vue suivantes (figure B.1.2). Cette approximation peut-être également mesurée à l'aide d'un appareillage ou être affinée à l'aide de données comme les points de fuite des bâtiments.

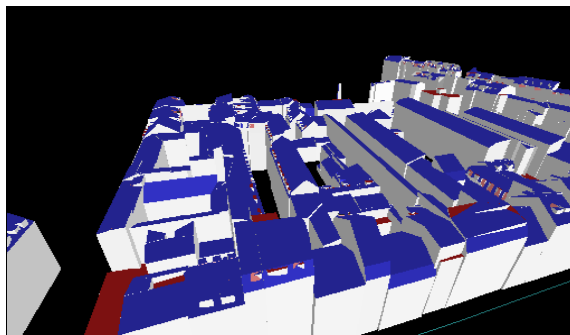


figure B.1.1 : vue perspective du modèle TRAPU utilisé



figure B.1.2 : clichés terrestres

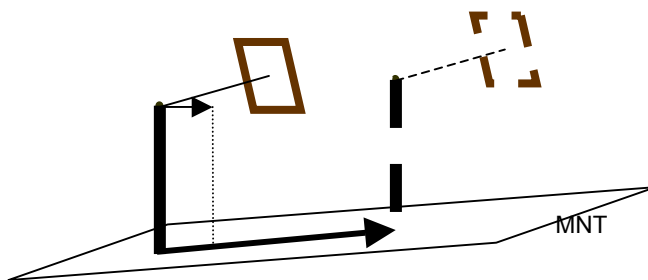


figure B.1.3 : clichés terrestres

Extraction de primitives

Les primitives utiles dans notre application sont les segments de droites et leurs extrémités, ainsi que les points de fuite principaux d'une image.

Segments de droites

Les algorithmes d'extraction de segments de droites proposés dans la littérature ne sont pas satisfaisants pour plusieurs raisons. Les extrémités des segments obtenus ne sont pas bien localisées et les algorithmes sont difficilement paramétrables. Nous avons développé un algorithme d'extraction de primitives linéaires connexes fiable ne nécessitant que deux paramètres [MARCHADIER, MICHELIN 99] [MARCHADIER et al 00].

L'originalité de ce travail réside dans la conservation d'une structure topologique (la carte de contours) obtenue à partir d'opérations bas niveau, et dans la formulation du problème d'extraction de segments de droites comme un problème d'optimisation d'un critère de sélection de modèles.

La carte de contours initiale peut être construite à l'aide d'algorithmes permettant l'obtention des crêtes du module du gradient de l'image initiale. Plusieurs algorithmes de ligne de partage des eaux ont été étudiés. L'algorithme de Marc Pierrot Deseilligny a été retenu et utilisé pour obtenir la carte de contours initiale. Cet algorithme a été testé sur des images terrestres (figure B.1.3). Un seuillage par hystérésis des courbes maximales locales permet de simplifier le résultat (figure B.1.4).

Les segments de droites (ou, plus généralement, les segments de courbes) sont extraits du graphe ainsi obtenu en optimisant une fonction de sélection de modèle.

Le critère utilisé dans l'algorithme développé est le critère de longueur de codage minimum, largement étudié dans la littérature de vision par ordinateur. Celui-ci réalise un compromis entre l'adaptation des données au modèle, caractérisée par la probabilité conditionnelle des données connaissant un modèle et ses paramètres, et la longueur de description du modèle.

$$s(c_j, \theta) = -\log P(c_j / \theta) + \text{sizeof}(m_i).$$

Les paramètres optimaux suivant ce critère pour un modèle donné sont ceux qui réalisent le maximum de vraisemblance.

Les données associées à chaque pixel du graphe sont ses coordonnées, ainsi que la direction du gradient au pixel considéré. Cette dernière permet de favoriser les chaînes de contours des régions homogènes des images. Le modèle de segment de droite est constitué de deux équations par point, de la forme suivante :

$$f(x_p, y_p, \theta_p; \theta, d) = \begin{bmatrix} x_p \cos \theta + y_p \sin \theta - d \\ \theta - \theta_p \end{bmatrix} = \begin{bmatrix} -r_{d,p} \\ -r_{\theta,p} \end{bmatrix} = 0$$

En considérant que les résidus des deux équations suivent des lois normales, et après un rapide développement, nous obtenons la fonction de sélection de modèle suivant pour chaque droite :

$$s(c_j, \hat{\theta}, \hat{d}) \approx n_j \log \left(2\pi \frac{\sigma_\theta \sigma_d}{\varepsilon_\theta \varepsilon_d} \right) + \frac{1}{2 \ln 2} \left(\sum_P \left(\frac{r_{d,p}}{\sigma_d} \right)^2 + \sum_P \left(\frac{r_{\theta,p}}{\sigma_\theta} \right)^2 \right) + (2 + n_j) t_{real}$$

Cette fonction de coût peut être étendue à d'autre type de primitives, comme par exemple des segments d'ellipse.

Une contrainte supplémentaire permet d'ignorer les courbes ayant un petit coût, mais ne correspondant pas à une réalité image (figure B.1.3).

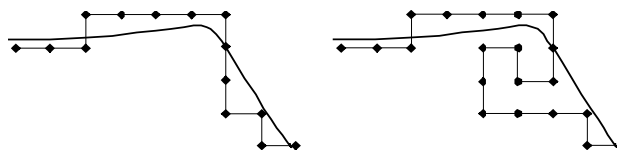


figure B.1.4 : à gauche, une courbe admissible, à droite, une courbe rejetée

Un modèle de bruit est inclus, de façon à écarter les points de contours qui ne sont pas significatifs.

La fonction minimisée est alors la somme des coûts des primitives (bruit ou segment de droites) réalisant une partition parfaite de la carte de contours. C'est un problème d'optimisation combinatoire qui ne peut être résolu qu'en examinant toutes les solutions possibles, ce qui est en pratique infaisable.

Nous avons développé un algorithme sous-optimal qui permet d'obtenir en un temps raisonnable une solution au problème. Cet algorithme est fondé sur une stratégie gloutonne assez rapide. Plusieurs critères de relaxation d'étiquette ont également été étudiés. Les deux seuls paramètres de l'algorithme sont les écarts types associés aux deux équations du modèle de segment de droite. Ceux-ci peuvent être estimés ou non, suivant le type de résultat souhaité.

Le résultat obtenu est satisfaisant (figures B.1.5, B.1.6 et B.1.7) et les extrémités de nombreux segments sont correctement localisées. Il est possible d'obtenir un algorithme fiable qui crée une segmentation d'une image en utilisant le même type de critère. Il faut pour cela prendre en compte des modèles de régions (texture, éclairage). Cet algorithme serait d'autant plus fiable qu'il utiliserait la dualité des modèles de régions et des modèles de contours. Faute de temps, nous n'avons pas pu étudier cet aspect.



figure B.1.5 : image initiale



figure B.1.6 : carte de contours

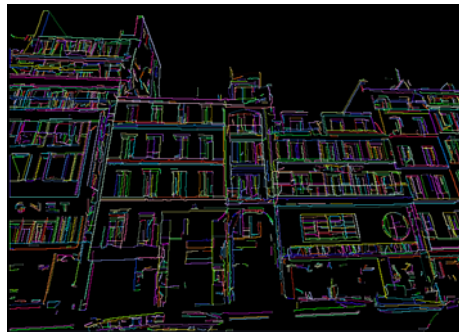


figure B.1.7 : segments de droites

Points de fuite

Nous avons développé une méthode permettant de calculer de façon précise la position des deux points de fuite principaux des images. Deux groupements de droites sont constitués. Chacun des groupes contient les droites qui participent éventuellement au faisceau de droite associé à un point de fuite. Une transformée de Hough peut éventuellement être utilisée pour déterminer les groupes initiaux de manière plus robuste. Dans notre application, nous avons utilisé une connaissance *a priori* sur le procédé de prise de vue. La position du point de fuite associé à chacun des groupes est déterminée à l'aide d'un estimateur robuste, qui sert à déterminer précisément la position du point de fuite et à éliminer les fautes.

Les équations utilisées lors de l'estimation d'un point de fuite sont de la forme :

$$\cos \theta_i (d_p, x_{ij} - \cos \theta_p) + \sin \theta_i (d_p, y_{ij} - \sin \theta_p) = 0$$

où (x_{ij}, y_{ij}) sont les coordonnées du point j appartenant au segment i , (θ_i, d_i) sont les paramètres de la droite support du segment i , et (θ_i, d_p) sont les paramètres (sous forme de coordonnées polaires, d_p étant l'inverse de la distance du point à l'origine) du sommet du faisceau (point de fuite).

Ce type de paramétrage nous permet de prendre en compte l'ensemble des points de fuite possibles.

La technique robuste utilisée est composée en deux étapes. La première étape est l'estimation à l'aide d'un estimateur par consensus. L'estimateur retenu est RANSAC. La seconde étape est une technique d'élimination de fautes par une méthode par bouclage.

À chaque itération de cet algorithme, un test du χ^2 pour chaque droite participant au faisceau, rejetant les droites échouant le test et acceptant les droites pour une nouvelle itération dans le cas contraire. Une nouvelle estimation est alors menée. L'algorithme s'arrête quand les droites sélectionnées sont stables entre deux itérations successives.

Appariements initiaux guidés par le modèle

Le problème d'appariement automatique de primitives entre images est un problème assez classique en vision par ordinateur. Les primitives utilisées peuvent être des points caractéristiques d'une image ou des primitives plus complexes (chaînes de contours, segments de droite...), suivant la problématique sous-jacente.

Trois grandes problématiques peuvent nécessiter un algorithme d'appariement :

- déterminer les paramètres de prise de vue d'une série d'images ;
- la reconstruction d'une scène ;
- l'indexation.

La fiabilité du processus d'appariement (qui peut être caractérisée par la proportion d'appariements corrects obtenus) est directement dépendante de la précision des paramètres de prise de vue. Ceux-ci ne sont déterminés précisément qu'avec un ensemble d'appariements fiables. Ce dernier point met en évidence la complexité de la première problématique décrite, celle qui nous intéresse.

Des connaissances supplémentaires sur le type d'image analysée peuvent donner des contraintes qui ont une double utilité : elles permettent de fiabiliser le processus d'appariement et de limiter la combinatoire de celui-ci. Les contraintes de nature structurelle peuvent être introduites à l'aide de différents types de graphes mettant en relation les primitives d'une image. Des contraintes géométriques sont aussi classiquement utilisées, comme la limitation de la profondeur de la scène ou la contrainte épipolaire.

Les contraintes de nature radiométrique (corrélation, mesures de similitude invariantes à des transformations géométriques) sont généralement considérées. Les procédés d'appariements non denses peuvent être regroupés en deux paradigmes, à savoir la mise en relation de graphes (isomorphisme...) et le paradigme hypothèse-vérification (techniques de clustering, estimation robuste du modèle géométrique décrit...). Nous proposons ici une stratégie d'appariement qui est relativement efficace (en temps de calcul) qui repose sur l'utilisation de contraintes fournies par le modèle tridimensionnel, le procédé de prise de vue, et l'environnement photographié.

Le modèle 3-D fournit des informations de nature différente :

- des informations intrinsèques : les primitives modélisées, leurs relations géométrique (angles entre deux segments de droites, droites parallèles...) et topologiques (adjacence de primitives) ;
- des informations extrinsèques qui peuvent être utilisées quand une approximation des paramètres de prise de vue est connue.

Nous supposons que les paramètres de prises de vue sont approximativement connus. Dans une application donnée, ceux-ci peuvent être trouvés avec un appareillage adéquat (gyroscope, GPS, capteur Polhemus), ou par prédiction de mouvement (connaissance *a priori* sur le procédé de prise de vue, extrapolation de mouvement).

Focalisation

Dans notre application, nous restreignons l'espace de recherche entre primitives (segments de droites et points) d'images différentes en créant un groupe de segments potentiellement homologues par image et par segment du modèle 3-D. Ceux-ci sont constitués en sélectionnant les segments détectés sur les images se trouvant à l'intérieur d'une zone autour de la projection de chaque segment visible du modèle (figure B.1.8). La taille de cette zone est déterminée en fonction de l'incertitude des paramètres de prises de vue.

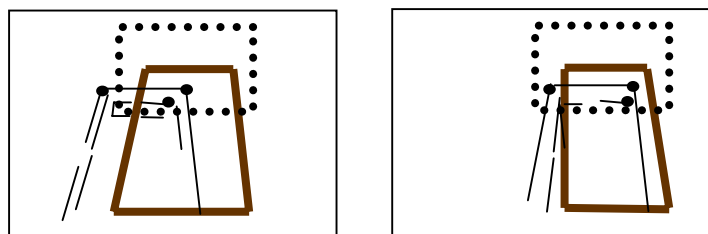


figure B.1.8 : restriction de l'espace de recherche

Contrainte de recouvrement

Un filtrage supplémentaire est effectué au sein de chaque groupe de segments homologues associé à un segment du modèle. Celui-ci élimine les segments images potentiellement homologues dont les « projections » sur la droite supportant le segment du modèle considéré sont disjointes.

La projection d'un segment d'une image sur une droite du modèle est obtenue en trouvant les intersections des plans perpendiculaires au plan formé par le centre optique d'une image et le segment du modèle considéré, et passant par le centre optique et chaque extrémité d'un segment de l'image, avec la droite supportant le segment modèle [ZHANG 96].

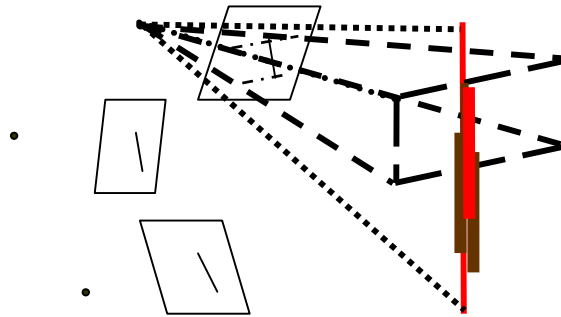


figure B.1.9 : contrainte de recouvrement

Appariements des extrémités

Les appariements de points utilisés dans la suite de cet article sont déduits des groupes de segments homologues. Les couples d'extrémité des segments retenus sont ceux réalisant le maximum de corrélation sur un petit voisinage. L'ensemble de ces opérations donne environ 60 % d'appariements corrects avec le meilleur jeu de paramètres. L'un des points faibles de la méthode est l'utilisation d'une corrélation qui s'avère peu robuste face à des déformations géométriques importantes.

Contrainte des points de fuite principaux

Un filtrage supplémentaire peut être réalisé à l'aide d'une approche photogrammétrique. Un estimateur robuste de la géométrie épipolaire pour un couple de clichés est classiquement utilisé. Nous avons conçu un nouvel outil de filtrage « géométrique » qui utilise les points de fuite des images.

Cette méthode fonctionne pour des points à peu près coplanaires. Elle tire partie de la connaissance de deux points de fuite issus de direction perpendiculaire sur chacune des images et de leur correspondance entre des images successives. Elle peut être décomposée en deux étapes distinctes, ce qui permet de simplifier énormément les différents calculs.

Pendant la première étape, la transformation qui permet de redresser une image est calculée, connaissant les deux points de fuite principaux. Puis, la transformation qui permet de superposer deux images consécutives est déterminée.

Redressement d'une image

Appelons P_1, P_2 : deux points de fuite principaux d'une image, exprimés en coordonnées homogènes. Il est possible de calculer l'homographie H qui envoie les points P_1, P_2 à l'infini $P_{1\infty}, P_{2\infty}$. Cette transformation s'exprime simplement à l'aide d'une application linéaire dans le plan projectif :

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & b & -f \end{bmatrix}$$

Les points de fuite obtenus ne sont pas nécessairement orthogonaux, ce qui est requis pour un éventuel redressement des images. Il est alors nécessaire de calculer le cisaillement C tel que $P_{1\infty} \perp P_{2\infty} = 0$. Celui-ci s'exprime de manière simple. À l'aide de la seule connaissance de deux points de fuite principaux, nous pouvons calculer un redressement simple, en composant les deux transformations précédentes.

Cette transformation s'écrit, toujours à l'aide d'une formulation projective : $H' = CH$. La figure B.1.10 (page suivante) montre l'effet de cette transformation (à droite) sur une image fortement déformée par la perspective (à gauche).



figure B.1.10 : redressement à l'aide des deux points de fuite principaux

Images consécutives

Deux clichés successifs d'une bande sont, une fois redressés, superposables à une translation et à une homothétie près. Plus formellement, deux points homologues P et P' de deux clichés successifs suivent la relation :

$$P = \begin{bmatrix} h_x & 0 & t_x \\ 0 & h_y & t_y \\ 0 & 0 & 1 \end{bmatrix} P'$$

Les quatre paramètres de cette contrainte sont estimés à l'aide de tous les homologues d'un couple. L'estimateur employé est RANSAC, qui peut déterminer les paramètres d'un modèle acceptable avec plus de 50 % de fautes d'observation. L'estimation est très efficace en temps de calcul, car le modèle estimé est très simple.

Les « fautes » détectées sont : soit, des appariements réellement aberrants, soit, des homologues qui ne sont pas dans le plan défini de manière approximative par les façades des bâtiments. À l'issue de ce filtrage, la totalité des appariements sont corrects. Ils peuvent être utilisés pour calculer une orientation relative très fiable.

Construction d'une mosaïque

En répétant le procédé décrit au paragraphe précédent sur chaque couple d'images successif, et en ramenant l'ensemble des images dans un repère associé au premier cliché, il est possible de construire une mosaïque d'un ensemble de façades d'une rue. La figure B.1.11, est constituée de trois clichés consécutifs.

Dans cet exemple, la superposition des images est réalisée à l'aide d'une simple moyenne des valeurs de chaque pixel, ce qui n'est évidemment pas satisfaisant, en particulier pour les problèmes d'occlusions.



figure B.1.11 : une mosaïque

Conclusion

Dans cet article, qui relate une partie des efforts de recherche que nous avons réalisés cette année, nous avons proposé quelques techniques qui permettent l'obtention d'appariements fiables entre images. Ces techniques n'ont pas fait encore l'objet de publication, bien qu'elles soient relativement innovantes.

Ces techniques, très spécifiques à notre problématique, sont efficaces et robustes. Leurs inclusions dans une chaîne de production complète peuvent être envisagées.

Référence

[GANG, ZHENGYOU 96] Gang Xu, Zhengyou Zhang : *Epipolar Geometry in Stereo, Motion and Object Recognition*. Kluwer Academic , volume 6 of Computational Imaging and Vision, Kluwer Academic Publishers, 1996

Publications

[MARCHADIER et al 00] J. Marchadier, S. Michelin & Y. Egels : *A graph labelling approach for connected linear feature selection IAPR Workshop on Syntactic and Structural Pattern Recognition, SSPR'2000*, in Lecture Notes in Computer Science, n° 1876, laboratoire MATIS, IGN-SR 00-015/COM-JM, 2000.

[MARCHADIER, MICHELIN 99] J. Marchadier & S. Michelin : *Extraction de primitives linéaires connexes d'une carte de contours*, AFIG'99, laboratoire MATIS, IGN-SR 990028/COM-JM, 1999.