

Méthodes de Correspondance d'Images Numériques

H. BOUNOUR, F. BOUCHENAK, A. DJILALI

Centre des Techniques Spatiales

BP, 13 Arzew-31200 – Algérie

bounour39@hotmail.com

ملخص :

يتمثل مبدأ الرؤية المجسمة في إستعمال عدة صور من نفس المنظر مأخوذة من مناظر مختلفة. يسمح هذا، بإستنتاج معلومات ثلاثية الأبعاد 3D على المنظر ابتداءً من إسقاطه الثنائي الأبعاد 2D على الصور. تطبيقات الرؤية المجسمة متنوعة : الملاحة الآلية ، تشكيل الخريطة و الإستكشاف الجوي، إعادة تشكيل العناصر، إلخ. غير أنه، يمكن تحديد موقع نقطة في الفضاء إذا كان موقع إسقاطها معروف في كل صورة. لا يوجد حل إجمالي لتطابق الصورة المجسمة بسبب وجود الغموض (الكسوف، إنعكاس السطوح، نقص النسيج، إلخ)، عموماً، لحل هذا المشكل، تمّ تفحص عدد من القواعد (على سبيل المثال Géométrie épi polaire) و الفرضيات (الإضاءة الإجمالية الثابتة، السطح المتماثل ، إلخ). تهدف المقالة الحالية إلى شرح بعض طرق تطابق الصور المجسمة التي تسمح بتخفيف مشاكل الغموض .

Résumé : Le principe de la stéréovision consiste à utiliser plusieurs images d'une même scène prises sous différents points de vue. Ceci, permet de déduire des informations 3D sur la scène à partir de ses projections 2D sur les images.

Les applications de la stéréovision sont diverses : navigation pour la robotique, construction de carte et reconnaissance aérienne, reconstruction d'objets, etc.

Néanmoins, la détermination de la position d'un point dans l'espace n'est possible que si la position de sa projection dans chaque image est connue. En raison, principalement des ambiguïtés (occultations, réflexions des surfaces, manque de textures, ... etc.), il n'existe pas de solution globale de mise en correspondance. En général, pour résoudre ce problème, un certain nombre de contraintes (par exemple la géométrie épipolaire) et d'hypothèses (illumination globale constante, surface uniforme, etc.) sont considérées.

Le présent article a pour objectif de présenter quelques méthodes de mise en correspondance d'images stéréoscopiques permettant de pallier le problème d'ambiguïtés.

Mots clés : stéréovision, projections, mise en correspondance, ambiguïtés, géométrie épipolaire.

Abstract : The principle of the stereovision consists in using several pictures of the same scene taken under various points of view. This allows to deduct information 3D about the scene from its projections 2D on the pictures. The applications of the stereovision are different: construction of map and aerial recognition, navigation for the robotics, reconstruction of objects, etc.

Nevertheless, the determination of the position of a point in the space is possible that if the position of its projection in every picture is known. In reason, mainly of ambiguousness appropriate in correspondence (eclipses, reflection of surfaces, lack of textures, ... etc.), it doesn't exist a global solution appropriate of image matching. In general, to solve this problem, a certain number of constraints (for example geometry épipolaire) and of hypotheses (constant global illumination, uniform surface, etc.), are considered. The present article has for objective to present some methods of image matching permitting to palliate problems of ambiguousness.

Keywords : stéréovision, projections, image matching, ambiguities, épipolar geometry.

1. Introduction

La photogrammétrie demeure une technique encore très compétitive, tant par sa précision que par sa fidélité, pour la production de l'information géographique. Cependant, cette technique nécessite l'emploi de matériel coûteux et d'un personnel qualifié pour l'interprétation, l'extraction sémantique et la restitution de l'information photographique.

Ces contraintes ont incité les chercheurs à tenter de rendre la photogrammétrie plus accessible, en essayant d'automatiser le processus photogramétrique. Ainsi, le développement en matière de micro-ordinateurs et de capteurs a rendu les problèmes photogramétriques plus solubles, et ceci par l'utilisation d'algorithmes exploitant des images numériques.

La photogrammétrie numérique, à l'instar des autres processus, est basée sur l'emploi strict d'images numériques et non de photographies conventionnelles.

Parmi les principaux problèmes que doit résoudre celle nouvelle approche, est l'identification des rayons homologues sur une paire d'images d'une manière automatique, connue sous le nom de "mise en correspondance ou corrélation d'images".

L'inconvénient majeur de la mise en correspondance, c'est quelle est sujette à certaines ambiguïtés telles que les occultations, les réflexions des surfaces, le manque de textures, etc. Pour pallier à ce problème, nous proposons à travers cet article quelques techniques de mise en correspondance d'images stéréoscopiques (numériques) qui sont généralement utilisées en aérotriangulation.

2. Notions générales

La mise en correspondance des points homologues pour la reconstitution de surfaces tridimensionnelles de l'espace objet est une opération délicate et fondamentale pour la vision stéréoscopique numérique. Elle consiste à retrouver dans des images à recouvrement, les primitives homologues qui sont la projection de la même entité de la scène. Ainsi, la mise en correspondance d'image serait d'associer à chaque point de la première image, ses correspondants dans les autres images.

Par primitive, nous entendons tout élément simple ou composé appartenant à une image, comme :

- Pixel : plus petit élément d'une image.
- Segment : défini par ses deux extrémités (nœuds) et par son orientation.
- Contour : ensemble connexe de pixels, appartenant à une même ligne généralement fermée, définissant des objets dans l'image.
- Région : zone de l'image dans laquelle, les pixels présentent les mêmes propriétés (niveau de gris, moyenne, variance, ...etc.).
- Point caractéristique : correspondant à un pixel particulier, dont il est extrait généralement des frontières des régions, ou à partir des contours.

Dans le cadre de ce travail, les primitives sont les pixels des images.

La base de la vision en stéréoscopie repose sur le fait qu'un objet de l'espace est enregistré à partir de deux points de vue différents. Cette différence en position a pour effet de causer l'apparition de parallaxes. C'est à partir de cette dernière et des éléments d'orientations externes que la position tridimensionnelle de l'objet pourra être calculée. Ainsi, tout algorithme de stéréovision est basé sur l'exécution chronologique des étapes suivantes :

- Détection des phénomènes
- Correspondance de ces phénomènes
- calcul 3D
- Interpolation pour la couverture de tout l'objet.

3. Mise en correspondance d'images

La mise en correspondance est caractérisée par ses attributs mis en œuvre et par la technique d'appariement (locale ou globale) qui est répartie en deux catégories [Hahn, 1995] :

- *Mise en correspondance locale* : Technique qui satisfait une mesure de ressemblance entre les primitives d'images.
- *Mise en correspondance globale* : Ensemble d'appariements qui satisfont à des contraintes globales.

3.1 Principes et définitions

Considérons deux images 1 et 2 d'une même scène. Il s'agit de déterminer, pour un élément de l'image 1, l'élément correspondant dans l'image 2. Cette mise en correspondance de primitives qui est un problème fondamental de la vision par ordinateur, est un processus intermédiaire entre les processus dits de hauts niveaux (reconstruction, reconnaissance) et de bas niveaux (extraction d'indices).

La mise en correspondance qui reste ambiguë nécessite l'introduction de contraintes dans la recherche de l'élément correspondant. Une recherche de correspondance sans contraintes pourra donner des résultats ambigus et par conséquent entraînera une mauvaise interprétation. Afin d'obtenir de bons résultats (interprétation), il est nécessaire d'introduire des contraintes dans la recherche de l'élément correspondant. Ce problème est appelé «le problème de la fausse cible» [Grimson, 1981] (fig. 1).

Les règles de la vision naturelle sont définies de la manière suivante :

- Un point donné sur la surface de l'objet n'a qu'une seule position dans l'espace.
- Les surfaces des objets sont en générale continues. Ceci impose des règles en stéréovision sur les parallaxes qui sont l'unicité et la continuité. L'utilisation de ces règles permet de réduire les ambiguïtés dans la mise en correspondance par:
- Des contraintes géométriques sur la position

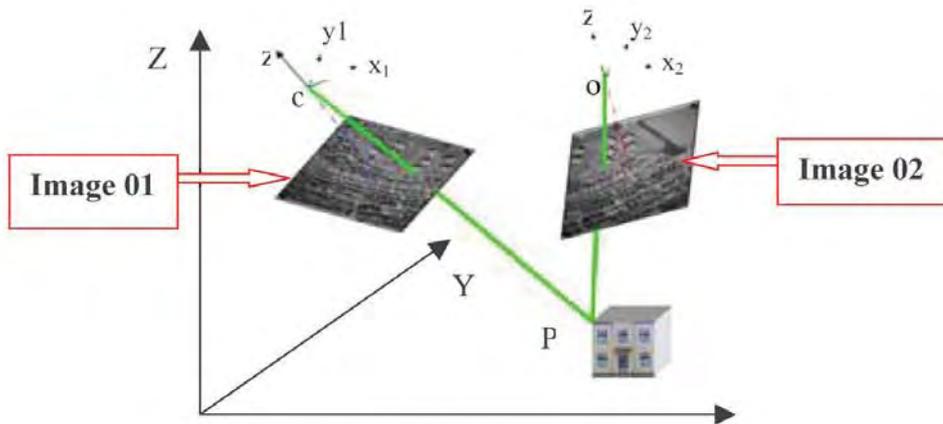


Fig.1 : Mise en correspondance

3.2 Problème d'ambiguïté dans la mise en correspondance

La mise en correspondance des images stéréoscopiques est très difficile à cause des points candidats multiples (fig. 2).

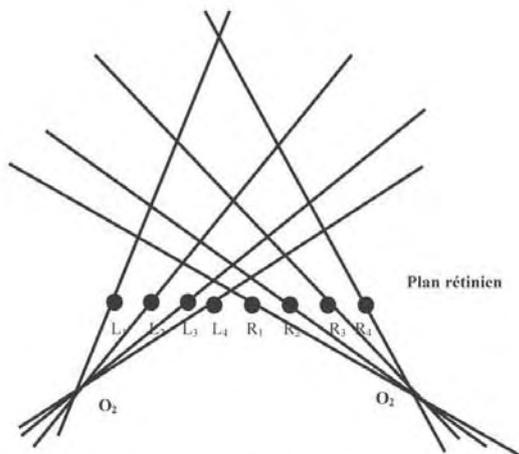


Fig. 2 Problème d'ambiguïté dans la mise en correspondance O_1O_2 : axe optique respectivement de la camera 1 et 2.

d'un correspondant dans l'image 2 sont utilisées pour réduire le nombre de correspondants potentiels. - L'introduction d'une mesure de similarité entre les primitifs dans deux images.

3.3 Contraintes de mise en correspondance

Afin de juger la validité de la mise en correspondance, certaines contraintes sont à vérifier ou à imposer.

3.3.1 Contrainte épipolaire

La contrainte épipolaire est une contrainte géométrique intrinsèque au capteur. Elle s'applique aux points et aux segments de droite. Elle réduit l'ensemble des correspondants potentiels d'un point à une droite dans l'image.

Considérons le point P1 de l'image 1 (fig.3). Les points de l'espace ayant pour image le point P1 sont situés sur la ligne de vue de direction d1.

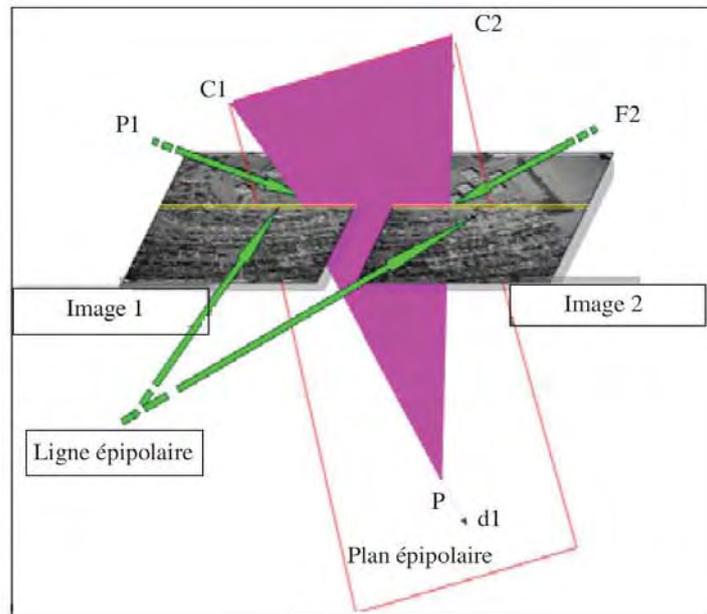


Fig. 3 Plan épipolaire.

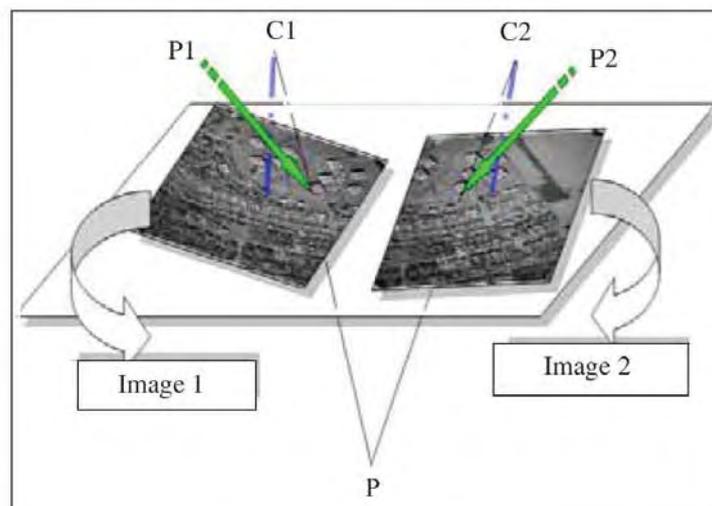
correspondants potentiels de P1 dans l'image 2 est appelée droite épipolaire de P1. Les droites épipolaires dans une image sont l'intersection entre l'image et le plan épipolaire.

3.3.1.1 Rectification d'images en géométrie épipolaire

La rectification d'images consiste à recalculer, pour deux images en position, deux nouvelles images telles que la géométrie épipolaire de ces deux images soit simple; c'est à dire que les droites épipolaires soient horizontales et parallèles à la base de prises de vues. Ceci, implique que les deux nouveaux épipoles soient à l'infini (fig. 4).

Une solution consiste à garder les deux centres de projection et à utiliser comme nouveaux plans rétinien un seul et même plan contenant la direction de la droite liant les deux centres de projection. Ce plan n'étant pas défini de manière unique, il faut choisir une orientation. On peut alors considérer que le nouveau plan rétinien contient la direction de la droite intersection des plans rétinien des images originales.

L'utilisation de la géométrie épipolaire s'applique soit pour un échantillonnage afin de redresser l'image, soit pour rechercher le déplacement de la fenêtre référence sur la fenêtre de recherche.



F ig. 4 Rectification d'images.

plan épipolaire de deux vues **1** et **2**. La contrainte d'ordre impose à ce que leurs projections **p** et **p0** apparaissent dans le même ordre chronologique sur les droites épipolaires des images **1** et **2**.

3.3.4 Contrainte d'unicité

Elle impose à ce qu'à tout point d'une image doit lui correspondre un et un seul point de l'autre image.

4. Techniques de la mise en correspondance

Les techniques de mises en correspondance peuvent être classées en deux grandes familles : les techniques *zonales* (area based) et *structurales* (features based).

4.1 Correspondance zonale (Area Based Matching ABM)

4.1.1 Définition

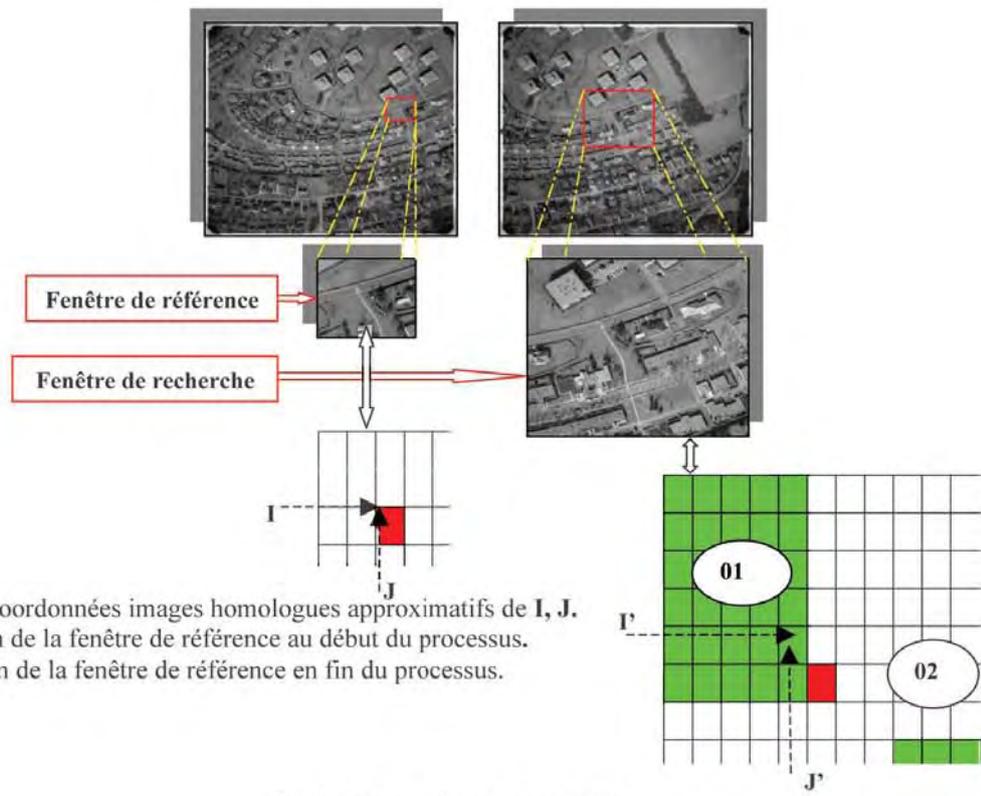
La correspondance zonale est la détermination des points homologues en se basant sur l'évaluation de la ressemblance de deux fenêtres à partir des images gauche et droite.

Cette évaluation est basée sur la comparaison des valeurs du niveau de gris des pixels dans chaque fenêtre respective, et cela en appliquant une analyse statistique ou une similitude spectrale. Le problème principal est de trouver la position géométrique sur une image numérique. Pour cela, on utilise dans un premier temps, la corrélation (cross-correlation) qui donnera une précision au niveau du pixel, et dans un deuxième temps, une corrélation par moindres carrés (least square corrélation) qui donnera une précision au niveau de sub-pixel [Rodriguez, 1992].

4.1.2 Coefficient de corrélation

Le processus de la mise en correspondance basée sur le calcul du coefficient de corrélation utilise des valeurs du niveau de gris des pixels de deux fenêtres appartenant chacune à une image, telle que :

- Choix d'une fenêtre de référence sur l'image de gauche centrée sur un pixel dont on cherche l'homologue.
- Calcul de la position de la fenêtre de recherche sur l'image de droite qui est d'une taille plus grande que celle de la référence (fig. 5).



(I', J') : les coordonnées images homologues approximatifs de I, J.
 (01) : Position de la fenêtre de référence au début du processus.
 (02) : position de la fenêtre de référence en fin du processus.

Fig. 5 Fenêtre de référence et de recherche

Le problème de ce choix se pose au niveau de la taille de la fenêtre qui dépend de la nature du terrain, de l'échelle de la photographie, de la résolution et de la texture.

Pour trouver la position de la fenêtre de référence dans l'image de recherche, on calcule le coefficient de la corrélation **R** donné par [Rodriguez, 1992]:

$$R = \frac{\sum_i \sum_j (g_{1ij} - m_1)(g_{2ij} - m_2)}{\sqrt{\sum_i \sum_j (g_{1ij} - m_1)^2 \sum_i \sum_j (g_{2ij} - m_2)^2}}$$

Ce coefficient est calculé pour toutes les positions possibles de la fenêtre de référence dans la zone de recherche. La position recherchée est celle correspondant à la plus grande valeur de **R**.

4.1.3 Mise en correspondance par moindres carrés

Le calcul de **R** permet de trouver la position du correspondant au pixel près. Pour atteindre une précision du subpixel, on utilisera la technique de mise en correspondance par moindres carrés. Le principe de ce processus, dans le cas où on suppose que l'image est normale ou on considère seulement la direction x, est comme suit : Soient **G₁(x)** et **G₂(x)** les valeurs vraies de niveau de gris représentant des quantités inconnues des fonctions images **g₁(x)** et **g₂(x)** qui sont respectivement perturbées par les bruits **n₁(x)** et **n₂(x)**, telles que [Ackermann, 1984] :

$$G_1 = g_1(x) + n_1(x) \tag{1}$$

$$G_2 = g_2(x) + n_2(x) \tag{2}$$

Supposons que **g₂(x)** est une transformation géométrique et radiométrique de **g₁(x)** :

$$g_2(x) = h_0 + h_1 \cdot g_1(a_0 + a_1 x) \tag{3}$$

- h₀ : correction du niveau zéro
- h₁ : facteur d'échelle de luminance
- a₀ : correction géométrique
- a₁ : facteur d'échelle

La linéarisation de la relation (03) donne :

$$(a_0, a_1, h_0, h_1) = (0, 1, 0, 1) \text{ où } g_2^0(x) = g_1(x) \tag{4}$$

On obtient :

$$g_2(x) = g_1(x) + g'_1(x)da_0 + xg''_1(x)da_1 + dh_0 + g'_1(x)dh_1 \tag{5}$$

$$\text{Avec : } g'_1(x) = dg_1(x)/dx \tag{6}$$

$$\Delta g(x) = G_2(x) - G_1(x) = g_2(x) - g_1(x) + n_2(x) - n_1(x) \tag{7}$$

En remplaçant **g₂(x)** par son expression dans l'équation (7), nous aurons :

$$\Delta g(x) - v(x) = g'_1(x)da_0 + xg''_1(x)da_1 + dh_0 + g'_1(x)dh_1 \tag{8}$$

$$v(x) = n_2(x) - n_1(x) \tag{9}$$

L'équation (08) représente, pour chaque pixel, une équation d'observation linéarisée qui relie les différences du niveau de gris observés **Δg(x)** avec les paramètres de transformation inconnus (**a₀, a₁ h₀ h₁**). La solution par moindres carrés minimise la différence résiduelle **v(x)**, et donne la fenêtre de corrélation optimum.

Dans le cas où on considère un repère bidimensionnel (on n'a pas de géométrie épipolaire), l'équation d'observation est la suivante :

$$g_2(x,y) - v(x,y) = g'_x(x,y).da_0 + x.g''_x(x,y)da_1 + y.g'_y(x,y)da_2 + g''_y(x,y)db_0 + x.g''_y(x,y).db_1 + x.g''_y(x,y).db_2 + dh_0 + g_1(x,y)dh_1 \tag{10}$$

$$\text{Avec : } g'_x = \frac{\partial g_1(x,y)}{\partial x} \quad g'_y = \frac{\partial g_1(x,y)}{\partial y}$$

En discrétisant la fonction **g**, les dérivées **gx'** et **gy'** s'écrivent :

$$g'_x = \frac{g_1(x+1,y) - g_1(x-1,y)}{2}$$

$$g'_y = \frac{g_1(x,y+1) - g_1(x,y-1)}{2}$$

$$g_2(x,y) - v(x,y) = g_1(x,y) + g'_x dx + g'_y dy + dh_0 + dh_1 g_1(x,y) \tag{11}$$

A partir de la correction géométrique, on a

$$\begin{cases} x = a_0 + a_1 x_0 \\ dx = da_0 + x_0 da_1 \end{cases} \quad \begin{cases} y = b_0 + b_1 y_0 \\ dy = db_0 + y_0 db_1 \end{cases}$$

Où **x₀** et **y₀** représentent les coordonnées obtenu par le coefficient de corrélation.

L'équation (11) devient :

$$g_2(x,y) - (g_1(x,y) + g'_x dx + g'_y dy + dh_0 + dh_1 g_1(x,y)) = v(x,y) \tag{12}$$

Où (**dx, dy, dh₀, dh₁**) sont les inconnues

$$\Delta g(x,y) = g_2(x,y) - g_1(x,y) = g'_x dx + g'_y dy + dh_0 + dh_1 g_1(x,y) + v(x,y) \tag{13}$$

Par exemple, pour une fenêtre de référence de taille 3x3, le système d'observation s'écrit :

$$\begin{bmatrix} g'_{x(1)} & g'_{y(1)} & 1 & g_{1(1)} \\ g'_{x(2)} & g'_{y(2)} & 1 & g_{1(2)} \\ g'_{x(2)} & g'_{y(2)} & 1 & g_{1(3)} \\ g'_{x(2)} & g'_{y(3)} & 1 & g_{1(4)} \\ g'_{x(2)} & g'_{y(4)} & 1 & g_{1(5)} \\ g'_{x(2)} & g'_{y(5)} & 1 & g_{1(6)} \\ g'_{x(2)} & g'_{y(6)} & 1 & g_{1(7)} \\ g'_{x(2)} & g'_{y(8)} & 1 & g_{1(8)} \\ g'_{x(9)} & g'_{y(9)} & 1 & g_{1(9)} \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dh_0 \\ dh_1 \end{bmatrix} = \begin{bmatrix} \Delta g_{(1)} \\ \Delta g_{(2)} \\ \Delta g_{(3)} \\ \Delta g_{(4)} \\ \Delta g_{(5)} \\ \Delta g_{(6)} \\ \Delta g_{(7)} \\ \Delta g_{(8)} \\ \Delta g_{(9)} \end{bmatrix} \tag{14}$$

La solution est estimée par moindres carrés et les coordonnées définitives sur l'image droite sont données par : $x = x_0 + dx$ et $y = y_0 + dy$.

L'image de droite est corrigée radio métriquement par rapport à l'image de gauche par h_0 et h_1 , avec :

$$h_0 = \sum \sum \Delta h_{0(i,j)} \text{ et } h_1 = \sum \sum \Delta h_{1(i,j)}$$

Image droite = $h_0 + h_1$ *Image gauche

4.1.4 Avantages et inconvénients de la correspondance zonale

Avantages [Greenfield, 1989]:

- Plus simple et moins coûteuse,
- Donne de bons résultats dans le cas des surfaces plates et des images ayant des propriétés photogrammétrie similaires,
- Permet d'avoir un grand degré d'invariance contre les distorsions géométriques de l'image.

Inconvénients :

- Exige un temps de calcul long,
- Taille de la fenêtre,
- Ne fournit pas d'information explicite ou sémantique.

4.2. Correspondences structural (Feature Based Matching: FBM)

Cette technique consiste à extraire séparément à partir de chaque image des éléments structuraux primaires en leur appliquant un critère de corrélation basé sur certaines propriétés particulières qui peuvent être résumées comme suit :

- L'emplacement du pixel ;
- L'orientation;
- L'angle entre deux segments consécutifs dans le polygone approximant le trait,
- La force (la magnitude du gradient, la pente du "zéro-crossing"),
- Les moments (le contour est traité comme une fonction de densité),
- Les paramètres de la forme (cercle, ellipse),
- La couleur (si elle est disponible),
- L'emplacement sémantique relatif.

Cette méthode ne fournit pas une bonne précision pour la photogrammétrie. Par contre, elle est généralement appliquée pour la vision artificielle. Deux types de structures sont considérés : le point et le contour. Chacun de ces éléments peut être apparié, ce qui engendre deux types de méthodes :

- Point : opérateur d'intérêt (Point Feature Matching)
- Ligne : correspondance basée sur les contours (Edge Based Matching).

5. Conclusion

En photogrammétrie, la mise en correspondances est utilisée pour divers champs d'applications, notamment pour le transfert de points d'une image à l'autre, le positionnement pour la mesure de points artificiels, l'exécution de l'orientation relative et le calcul de la position tridimensionnelle de points dans l'espace de l'objet.

Généralement, en aérotriangulation, c'est la technique de mise en correspondance zonale (ABM) qui est adoptée car elle donne une très bonne précision et elle est invariante par rapport aux distorsions d'images. En outre, elle permet une analyse statistique directe [Heipke, 1992].

Si le produit photogramétrique recherché est autre que l'information tridimensionnelle (par exemple une restitution planimétrique automatique), différentes techniques peuvent être employées selon une combinaison appropriée. Les résultats de l'application d'une des techniques servent comme éléments d'entrée pour d'autres techniques [Heipke, 1992].

Il existe différentes approches pour la mise en correspondance d'images numériques. Les techniques présentées dans cet article sont celles les plus utilisées en photogrammétrie aérienne basée sur la surface.

La méthode de mise en correspondance zonale (ABM) est la plus utilisée car elle s'adapte le mieux aux travaux de photogrammétrie et plus particulièrement la photogrammétrie aérienne. Elle s'adapte aussi avec les techniques d'estimations par moindres carrés et différents modèles mathématiques.

L'échantillonnage en géométrie épipolaire permet de réduire le domaine de recherche des homologues; cette technique peut faire l'objet d'une étude approfondie sur le redressement en géométrie épipolaire.

Références Bibliographiques

Ackerman F., 1984 : "Digital image correlation: Performance and potential application in photogrammetry". "Photogrammetric Record Universidad Stuttgart Vol 11 pp 429-439.

Greenfield, 1989 : "Experiments with edge based stereo matching". Photogrammetric Engineering and Remote Sensing. 55 (12): 1771-1777.

Grimson W.E.L., 1981 : "A computer implementation of a theory of human stereo vision," Philos. Trans. R. soc. London, Ser. B, Vol. 292, pp. 217-253.

Hahn M., 1995 : Navigation using Image Sequence, in: k.Linkwitz & U.Hanleiter, Eds., "High Precision Navigation 95", Dummler Verlag, Bonn, pages 464-470.

Heipke C., 1992 : "A global approach for least-squares image matching and surface reconstruction in object space", vol.58, n°3 (49 ref).

Heipke C., 1997 : "Automation of Interior, Relative, and Absolute Orientation". ISPRS Journal of photogrammetry and Remote Sensing, 52(1): 1-20.

Rodriguez G., 1992 : "Digital image processing". Book Addison- weskey publishing company 716p.