

Chapitre 2

Techniques de restitution d'images numériques

Ce chapitre traite des techniques d'extraction de données 3D à partir des images. Quelques considérations exposées ici ne sont pas spécifiques à la photogrammétrie numérique, comme c'est le cas par exemple pour l'aérotriangulation qui présente quelques aspects nouveaux pour des images numériques (l'extraction automatique de points de liaison par exemple). Nous commencerons par une synthèse portant sur les techniques d'amélioration d'image avec le point de vue du photogrammètre, c'est-à-dire quand la géométrie est critique (*cf.* section 2.1). Avec la même logique, les techniques de compression de données seront analysées pour le cas spécifique d'images numériques (*cf.* section 2.2). Puis un résumé au sujet de l'usage de GPS dans les avions sera présenté (*cf.* section 2.3) : ceci n'est pas bien entendu spécifique à la photogrammétrie numérique, et intervient dans beaucoup de problèmes pratiques de la photogrammétrie d'aujourd'hui. Puis l'extraction automatique de points de liaison sera vue sous le titre « automatisation de l'aérotriangulation », ainsi que son usage pour calculer un tableau d'assemblage automatique (*cf.* section 2.4), ce qui constitue la première étape du traitement d'un ensemble d'images. Puis nous concluons ce chapitre avec une étude des techniques utilisées pour les postes de travail de photogrammétrie numérique (*cf.* section 2.5). À la date d'achèvement de ce livre, les méthodes d'extraction automatiques disponibles pour les aspects planimétriques (la végétation, les bâtiments, les routes...) paraissent encore très incertaines et restent au niveau des laboratoires : elles ne peuvent pas encore être considérées comme opérationnelles, nous ne les avons donc pas présentées. Dans le chapitre 3, l'extraction automatique de l'altitude sera bien sur expliquée, avec tous les problèmes de MNT, MNS, MNE : c'est le seul processus opérationnel disponible utilisant toutes les possibilités offertes par les images numériques (c'est par

ailleurs le plus simple, car toute autre extraction de détails exige beaucoup plus d'intelligence humaine et d'interprétation de la part de l'opérateur, et il est donc aussi beaucoup plus difficile à mettre en œuvre sur un ordinateur).

2.1. Améliorations des images numériques¹

2.1.1. Manipulation d'histogramme

Le rehaussement du contraste d'image est un processus qui tend à rendre les images plus belles à utiliser, sur des critères parfois subjectifs. Sans pouvoir forcément dire à quoi devrait ressembler une bonne image, il est souvent possible de dire si elle est optimisée, suffisamment par exemple pour voir certains détails. Il existe donc plusieurs techniques d'amélioration d'image, dont l'utilisation des analyses en composantes principales, mais les manipulations d'histogramme figurent sans doute parmi les méthodes les plus employées.

2.1.1.1. Définition de l'histogramme d'une image

En statistiques, l'histogramme est la représentation bidimensionnelle de la fonction h de R dans R , appelée densité de fréquence, définie par :

$$\begin{aligned} h(x) &= h_i \text{ si } x \in [a_{i-1}, a_i[\\ h(x) &= 0 \text{ si } x < a_0 \text{ ou si } x \geq a_k \end{aligned} \quad \text{vérifiant la propriété } \int_{-\infty}^{+\infty} h(x)dx = 1 \quad [2.1]$$

En traitement d'image, l'histogramme est donc le diagramme représentant la distribution statistique (h_i, n_i) où h_i représente le i^{e} niveau de teinte possible parmi les N valeurs possibles et n_i le nombre de pixels présentant la valeur h_i dans l'image numérique. Par abus de langage, la fonction h est souvent appelée histogramme. Cette fonction est toujours normalisée de façon à pouvoir être utilisée comme une loi exprimant la probabilité qu'un niveau de teinte n_i d'être présent sur l'ensemble de l'image. Le niveau de gris est donc considéré comme une variable aléatoire qui prend ses valeurs dans un ensemble de taille N valant souvent 256 en raison du codage habituel des informations élémentaires de chaque canal sur un octet. Les nouveaux capteurs numériques, dont la dynamique intrinsèque dépasse plusieurs milliers de niveaux, nécessitent plus d'espace de stockage.

A priori, aucune hypothèse ne peut être faite sur la forme d'un histogramme et il est tout à fait normal d'avoir un histogramme multimodal, présentant plusieurs pics. Un histogramme correspondant à une distribution gaussienne est

1. Cette section a été rédigée par Alain DUPERET.

possible mais plus courant dans le cadre d'images satellitaires. La seule certitude concerne l'étendue, dont il convient qu'elle soit aussi grande que possible, jusqu'à éventuellement occuper les 256 niveaux disponibles.

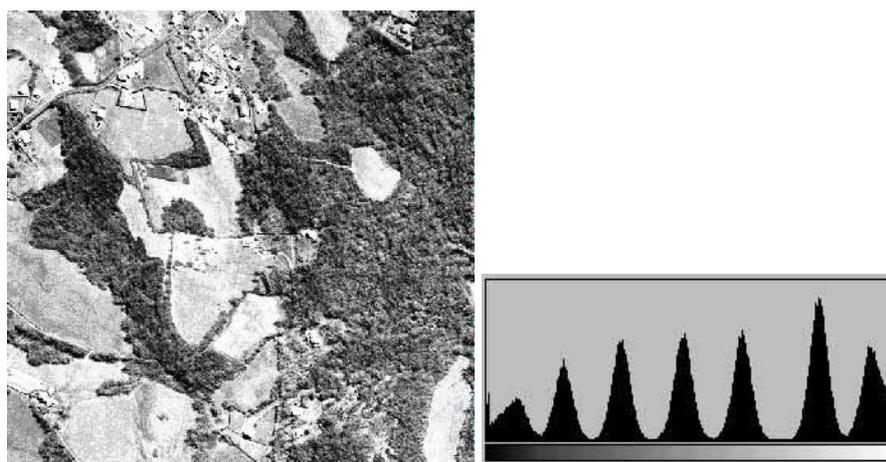


Figure 2.1. Extrait d'image aérienne – Histogramme associé

2.1.1.2. Représentation visuelle des données

Dans la pratique, lorsque les données présentent une dynamique faible pouvant nuire à leur exploitation, des transformations destinées à améliorer le contraste sont possibles et vont être présentées. Celles-ci agiront soit en modifiant les valeurs codées de l'image, soit en adaptant la table de transcodage utilisée par l'écran pour l'affichage de la couleur (LUT, *Look Up Table*). L'utilisateur devra également garder à l'esprit que lorsqu'une image est affichée sur un écran d'ordinateur, son contraste dépend également des réglages du moniteur mais aussi et surtout de la sensibilité spectrale de l'œil dont on sait qu'elle présente un maximum pour les teintes vert-jaune ; pour une variation de longueur d'onde $d\lambda$ donnée, cela entraîne une perception des teintes plus faibles pour le rouge ou le bleu-vert, situées aux bornes de la courbe de sensibilité spectrale de l'œil, elle-même variable d'un observateur à l'autre. D'autres facteurs sont également à prendre en compte, comme l'observation prolongée d'une couleur intense provoquant une aberration colorée transitoire due aux pigments rétiniens, l'environnement global du point observé qui peut faire apparaître une teinte plus ou moins claire, ainsi que les effets de bord.

2.1.1.3. Généralités sur la modification d'histogramme

Les manipulations d'histogrammes sont des opérations ponctuelles qui transforment donc une image d'entrée $I(x, y)$ en une image de sortie $I'(x, y)$ selon une loi $I'(x, y) = f[I(x, y)]$. La fonction f se représente sous forme d'un graphe pour lequel les abscisses sont les niveaux de teintes de I , et les ordonnées les niveaux de teintes de I' . Elles sont souvent destinées à recadrer la dynamique d'images en entrée initialement trop peu contrastées, trop claires ou trop foncées. Cette redistribution des teintes de l'image peut conduire à certains défauts d'interprétation lorsque la fonction f est non monotone ou discontinue :

- des paliers horizontaux dans le graphe de f conduisent à des pertes d'informations ;
- des sauts verticaux provoquent des discontinuités se traduisant par des faux contours ou des points isolés aberrants ;
- des pentes négatives conduisent à des inversions locales de contraste.

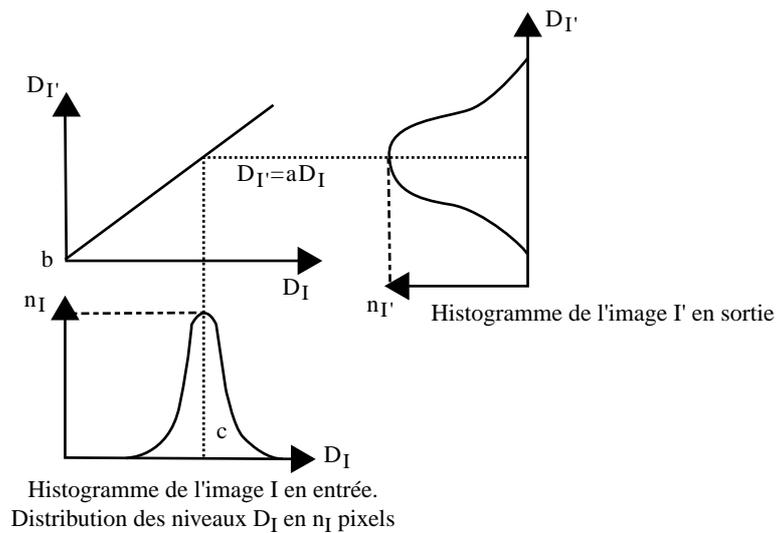


Figure 2.2. Manipulations d'histogramme

L'histogramme de l'image I' est centré en $ac + b$ et de largeur multipliée par a . Dans les logiciels de traitement d'images, le paramètre Luminosité peut être associé à la valeur b et le paramètre Contraste à la valeur a , pente de la courbe transformant D_i en $D_{i'}$.

2.1.1.4. Étalement de la dynamique

Comme l'image brute, considérée ici avec des pixels codés sur 1 octet, ne prend jamais toutes les valeurs possibles entre 0 et 255, cela se traduit par une carence de contraste, même si l'on sait que l'œil ne peut percevoir que quelques dizaines de niveaux de gris sur un écran. Il convient donc souvent d'étaler la dynamique initiale $[m, M]$ au plus grand intervalle possible $[0, 255]$ à l'aide de l'opération $D_I = (D_I - m) \times \frac{255}{M - m}$.

Cette opération ne se traduit par aucun apport d'information et relève plus de la stratégie d'affichage à l'écran que de la technique d'interpolation. Elle est souvent proposée sous une forme plus élaborée donnant de meilleurs résultats en calculant deux seuils min et max telles que $\int_0^{\min} h(x)dx = \int_{\max}^{255} h'(x)dx = \varepsilon$, $\varepsilon = 0,01$ ou $0,00$ (h représente la fonction de répartition des niveaux de teintes dans l'image). L'étalement pratiqué va alors de $[\min, \max]$ à $[0, 255]$.

Une variante intéressante existe lorsque l'histogramme en entrée est unimodal et relativement symétrique. On détermine alors min et max en fonction des propriétés statistiques de la distribution des teintes, comme l'illustre la figure 2.3.

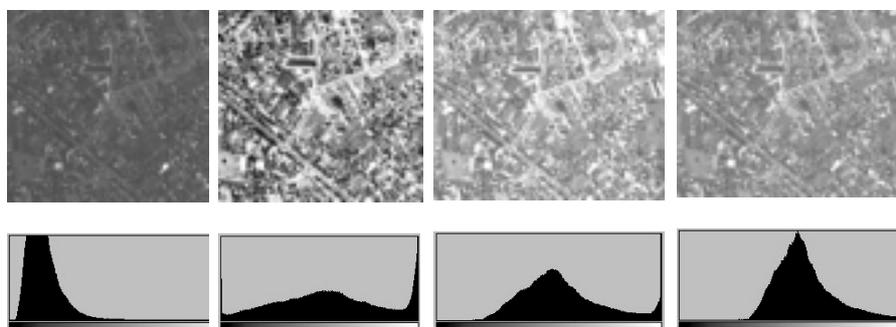


Figure 2.3. Pour une image brute donnée à gauche, les résultats visuels sont donnés pour un intervalle $[\min, \max]$ valant successivement $[\text{moy.} - \sigma, \text{moy.} + \sigma]$, $[\text{moy.} - 2\sigma, \text{moy.} + 2\sigma]$, $[\text{moy.} - 3\sigma, \text{moy.} + 3\sigma]$ en allant vers la droite

Pour une image multispectrale, l'étalement est réalisé indépendamment sur les différents canaux, au risque de provoquer une modification de la balance des couleurs, chacun des canaux ayant initialement une dynamique qui lui est propre pour une image brute donnée.

Il est tout à fait possible de procéder à un rééclairement non linéaire de la dynamique permettant de renforcer les contrastes pour une partie des teintes d'un histogramme : forme logarithmique pour réhausser les zones sombres, exponentielle pour les zones claires. Souvent utilisée, cette stratégie a le défaut de diminuer les contrastes sur les zones de l'histogramme non concernées par la forte courbure de la fonction de transfert mais elle permet de reproduire plus fidèlement la manière dont l'œil répond aux différents niveaux de luminosité.

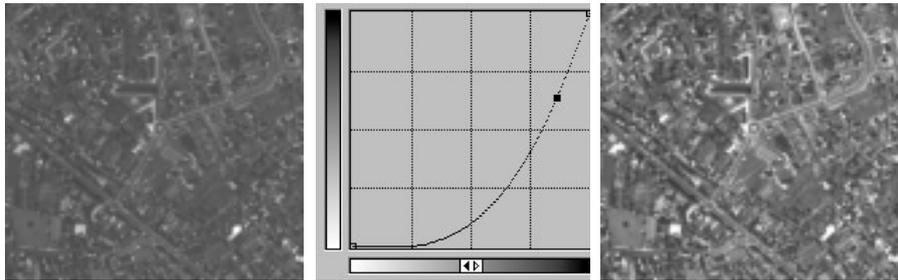


Figure 2.4. *Dispersion non linéaire de la dynamique. À gauche l'image originale, à droite l'image modifiée, au centre la courbe utilisée*

Une forte pente sur la courbe centrale traduit une accentuation des contrastes dans les zones correspondantes, qui sont ici les nombreuses parties sombres de l'image de droite. L'écrasement des contrastes dans les parties claires n'est ici pas visuellement gênant.

2.1.1.5. Égalisation d'histogramme

L'égalisation d'histogramme est une méthode qui recherche une distribution statistique équiprobable des niveaux h_r , ce qui revient donc à poser $p_s(s) = c$ où c est une constante.

La relation $p_r(r) dr = p_s(s) ds$ étant valable pour toutes valeurs, cela conduit donc à $\int_0^s cds = \int_0^r p_r(r) dr \Rightarrow c = \frac{1}{s} \int_0^r p_r(x) dx$ qui fournit la relation recherchée.

La précédente formulation présume que r et s peuvent prendre des valeurs continues. Avec les images numériques, il ne peut y avoir un nombre infini de valeurs dans l'intervalle $[r, r + dr]$. Lors de la redistribution sur un domaine plus large de cet intervalle, le principe de la conservation du nombre de pixels fait que seuls les pixels concernés sont redistribués et que l'histogramme ne peut donc être véritablement plat.

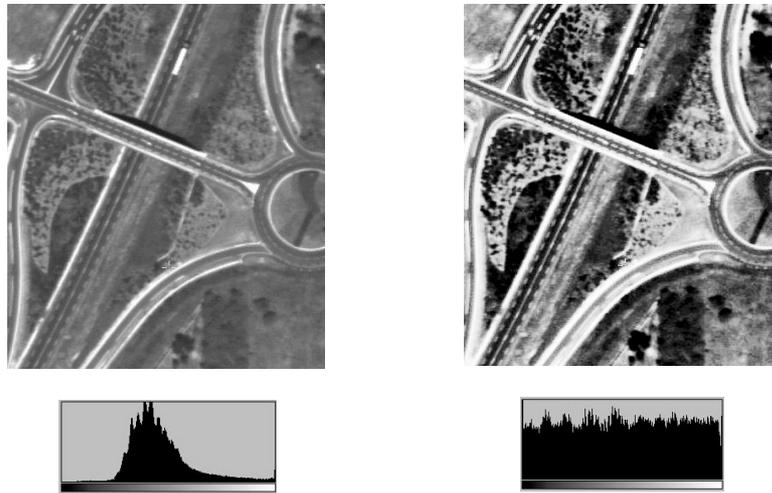


Figure 2.5. À gauche, image originale et son histogramme. À droite, image après égalisation avec son histogramme

Pour une image I constituée de n niveaux de teintes quantifiés sur N niveaux $D_I(k)$, ces derniers sont transformés par la loi :

$$D_I'(k) = (N - 1) \sum p_I(D_I(k)) \quad \text{où} \quad p_I = \frac{n_I}{n},$$

n_i représentant le nombre de pixels de valeurs $D_I(k)$. Les résultats $D_I'(k)$ pouvant être décimaux, la méthode implémentée dans le logiciel arrondit ensuite à la valeur entière la plus proche, ce qui conduit ainsi à avoir plusieurs niveaux $D_I(k)$ associés à $D_I'(k)$ qui se voit alors affecté la somme des probabilités des $D_I(k)$ dont il est issu : $p_I'(D_I'(k)) = \sum_i p_i(D_I(k))$ justifiant

ainsi la non-planéité de l'histogramme. Pour pallier cet effet, certaines méthodes redistribuent les pixels dans les teintes les plus représentées vers les valeurs voisines insuffisamment représentées, jusqu'à obtenir une population équilibrée pour chaque niveau de quantification.

Le nombre de teintes dans I' est donc inférieur à celui de I , tous les niveaux possibles n'étant pas forcément remplis. Cette technique permet d'obtenir la meilleure dynamique possible et procure des contrastes forts. Néanmoins, elle ne donne pas forcément de bons résultats visuellement, les teintes des objets photographiés ne se prêtant pas forcément à une telle hypothèse.

2.1.1.6. *Obtention d'un histogramme de forme quelconque*

L'opérateur peut très bien choisir d'avoir la fonction $p_s(s)$ comme histogramme final, ce qui revient alors à $\int_0^s p_s(y)dy = \int_0^r p_r(x)dx$, qui permet donc de définir s en fonction de r .

Dans la pratique, les étapes suivantes sont réalisées :

- égalisation de l'image originale $y_1 = g(r) = \int_0^r p_r(u)du$;
- égalisation de l'histogramme souhaité $y_2 = f(s) = \int_0^s p_s(v)dv$;
- application de f^{-1} (fonction qui se calcule facilement par tabulation) à l'histogramme $g(r)$ soit $f(s) = g(r) \Rightarrow s = f^{-1}g(r)$ (les deux histogrammes étant identiques).

2.1.1.7. *Amélioration de contraste local*

Les méthodes précédentes sont des manipulations globales de l'image et ne prennent pas en compte les éventuelles disparités de contraste ou de qualité qui peuvent exister au sein d'une même image. Des techniques d'amélioration locale de contraste ont été développées pour éviter cet inconvénient. Sur un voisinage de chaque pixel de taille $n \times m$ qui va progressivement être déplacé sur tous les pixels de l'image, l'histogramme local est calculé afin de pouvoir ensuite appliquer les méthodes vues précédemment comme l'égalisation ou l'attribution d'une forme d'histogramme donnée (exemple : forme gaussienne). Comme, à chaque fois, seule une nouvelle colonne ou une nouvelle ligne apparaît avec la translation d'un pixel dans l'image, il est possible de ne pas répéter le calcul complet de l'histogramme en s'appuyant sur les calculs de l'itération précédente.

Ainsi, pour une image donnée, l'attribution définitive d'un niveau de teinte pour un pixel passe à chaque fois par un traitement sur une manipulation de l'histogramme des valeurs présentes dans un voisinage de taille donnée.

Le procédé statistique opère sur des sous-images de taille variable et va utiliser l'intensité moyenne et la variance, qui sont deux propriétés pertinentes pour modéliser l'apparence de l'image. Pour une valeur radiométrique moyenne $moy(i, j)$ à l'intérieur d'une fenêtre moyenne centrée sur le pixel de coordonnées (i, j) , et une variance radiométrique de $\sigma(i, j)$ à l'intérieur du même voisinage, la valeur de la radiométrie dans I peut être modifiée dans I' :

$$I'(i, j) = A \times [I(i, j) - moy(i, j)] + moy(i, j) \text{ avec}$$

$$A = \frac{kM}{\sigma(i, j)}, \text{ scalaire appelé facteur de gain} \quad [2.2]$$

M : valeur moyenne des radiométries sur toute l' image et $k \in]0, 1[$

de telle manière à rehausser le contraste localement pour les zones faiblement contrastées, A étant inversement proportionnel à $\sigma(i, j)$. La taille de la fenêtre de traitement est un critère prépondérant :

– faible (3 x 3 à 9 x 9), elle permet de rehausser le plus fortement possible le contraste local, avec pour effet corollaire de faire apparaître le bruit d'image (cf. figure 2.6 a) ;

– forte (50 x 50, 100 x 100...), elle permet de moyenner les valeurs de teintes sur de grandes zones, diminuant ainsi les disparités de luminosité aux différents endroits de l'image, mais ses effets locaux sont alors moins sensibles (cf. figure 2.6 b).

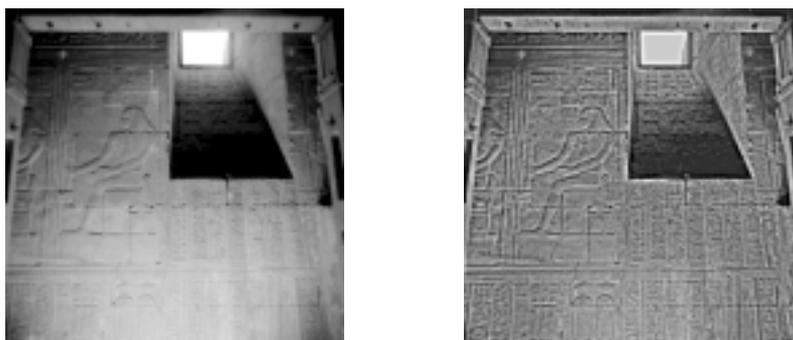


Figure. 2.6 a. Temple égyptien de Dendera – En photogrammétrie terrestre, les conditions d'éclairage sont parfois très médiocres. Les effets de sur ou sous-exposition sont diminués par une amélioration locale de contraste. Au-delà d'un certain seuil, le bruit d'image devient prépondérant

Ces méthodes sont très employées pour rendre semblables deux images prises à des dates différentes ; en imagerie aérienne, cela permet de réduire les disparités de radiométrie entre deux images distinctes lorsque celles-ci doivent être fusionnées pour faire une mosaïque, voire une ortho-image.

Au-delà du simple traitement d'une image isolée, la mise en œuvre sur un couple d'images stéréoscopiques peut se révéler d'un grand secours aux moyennes ou grandes échelles, tout particulièrement lorsque des zones d'ombres gênent la perception des objets.

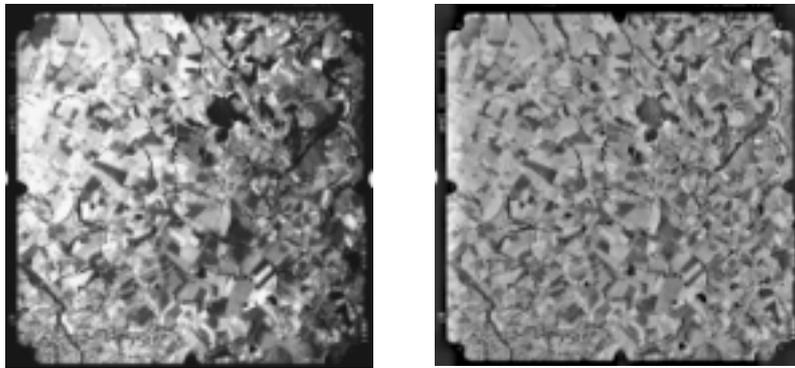


Figure 2.6 b. Zone d'Agen – L'image de gauche présente des zones claires dans le coin en haut à gauche (hot spot) qui peuvent être supprimées en imposant à l'image en sortie de droite de présenter un histogramme gaussien à l'intérieur d'une fenêtre de taille choisie par un opérateur

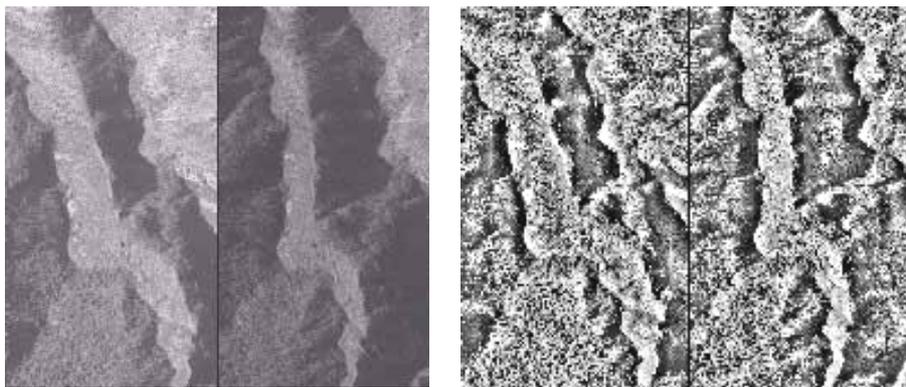


Figure 2.7. Zone de Roquesteron – Stéréogramme des images brutes à droite, des images avec rehaussement local de contraste sur des fenêtres de petite taille ; même si le bruit rehaussé apparaît rédhibitoire sur une image, l'examen stéréoscopique dans les zones d'ombre est malgré tout possible et permet ainsi de limiter le recours à un nouveau survol aérien ou à l'exploitation d'autres images

L'application à des images multispectrales (plusieurs canaux) est plus délicate, les traitements vus précédemment devant être appliqués à chacun des canaux indépendamment, tout en fournissant un résultat cohérent. Dans le cas d'images en couleur, la représentation peut se faire dans un référentiel colorimétrique en fonction de trois composantes principales rouge, vert et bleu ; chaque coordonnée d'un pixel le long de l'un de ces axes est le niveau de teinte

qu'il a dans la bande spectrale correspondante. L'ensemble des points correspondant aux pixels dans ce référentiel forme un nuage dont il convient alors de rechercher les axes le long desquels les données sont le mieux réparties. Cette recherche s'effectue à l'aide de la méthode de classification d'analyse en composantes principales couramment utilisée en télédétection.

L'opération de rehaussement local de contraste peut être réalisée en passant dans le domaine fréquentiel et en utilisant des techniques de filtrage dit homomorphique.

2.1.2. Filtrage – Suppression du bruit

2.1.2.1. Le bruit dans les images

Le signal contenu dans l'image est composé d'information, superposée avec du bruit ; dans le cas d'image, il peut s'agir du bruit inhérent au capteur numérique utilisé lors de la saisie, de la nature du document photographique s'il s'agit d'un document qui a été numérisé. Les méthodes vues précédemment traitaient chaque pixel indépendamment de la valeur des pixels voisins et ont aussi pour effet de rehausser sans faire de distinction le bruit et l'information dans les images, ce qui justifie la mise en place de traitements destinés à diminuer le bruit, visible sous forme de pixels ou de petits groupes de pixels isolés présentant de fortes variations avec leur environnement.

Dans la pratique, les images possèdent une certaine redondance spatiale, des pixels voisins ayant souvent des teintes proches. En effectuant des hypothèses générales sur le bruit (additif, indépendance vis-à-vis du signal et fortement décorrélé), il est possible de réaliser des opérations de filtrage destiné à limiter la présence du bruit dans les images en réalisant un filtrage passe-bas, linéaire ou non, dans le domaine spatial ou fréquentiel. La justification d'un tel filtrage est la suivante :

Le bruit dans les images est habituellement décorrélé du niveau du signal, et présente donc une distribution spectrale à peu près uniforme. Il se trouve que, pour les images, la part de basses fréquences est plus importante que celle des hautes fréquences. Il arrive donc un moment où le signal est dominé par le bruit, d'où l'idée de seuiller au-delà d'une fréquence de coupure. L'élimination de cette partie du spectre détruit bien entendu une certaine part du signal, ce qui se traduit par une image avec peu de texture et des contours plus flous.

Ce traitement peut s'appliquer par calcul de la Transformée de Fourier de l'image sur laquelle est appliqué un traitement retranchant toutes les fréquences au-delà de la valeur de coupure, puis par transformée de Fourier inverse.

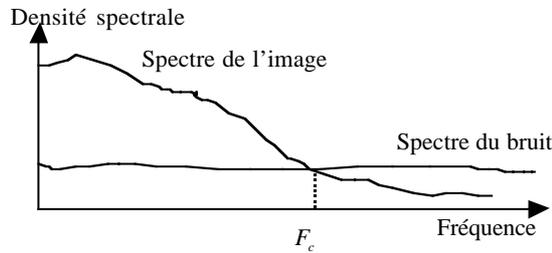


Figure 2.8. Au-delà de F_c , le spectre de l'image bruitée est dominé par le bruit

Il y a deux grands types de bruit :

- le bruit additif :
 - de type impulsionnel, dont la particularité est d'altérer aléatoirement les teintes de quelques pixels pour produire un effet « poivre et sel » dans les images ;
 - de type Gaussien, si une variation de teinte issue d'une densité de probabilité gaussienne centrée est rajoutée sur les teintes de l'image. C'est le type le plus courant ;
- le bruit multiplicatif, comme celui résultant d'une variation d'illumination dans les images.

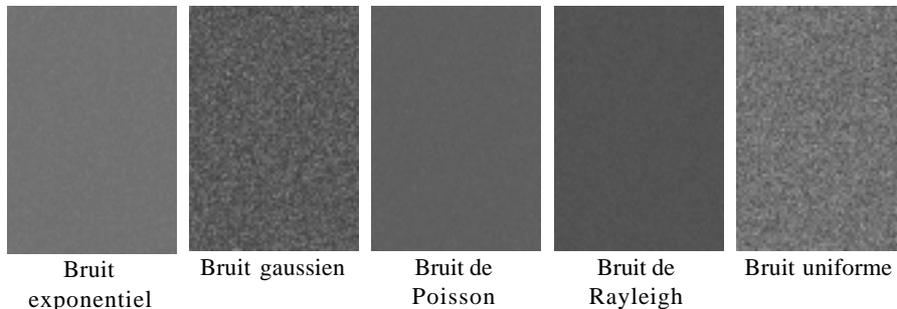


Figure 2.9. Représentation typique de différentes lois de bruit dans une image parfaitement uniforme

2.1.2.2. Mise en œuvre du filtrage par convolution

Les méthodes de filtrage sont nombreuses et se classent le plus souvent par type : filtrage linéaire, non linéaire, morphologique (ouverture, fermeture), par équation de diffusion (isotrope ou non), adaptatif par coefficients ou fenêtres adaptatives.

L'emploi d'opérateurs de convolution dans le domaine spatial (l'image) transforme une image $I(x, y)$ en $I'(x, y)$ à l'aide d'une fenêtre carrée de dimension impaire $2n + 1$ déplacée sur l'image, par une matrice A de coefficients a_{ij} de façon à ce que :

$$I'(x, y) = \sum_{-n \leq i \leq n, -n \leq j \leq n} a_{i,j} I(x + i, y + j) \quad [2.3]$$

2.1.2.3. Filtrage linéaire dans le domaine spatial

Le filtrage le plus simple consiste à effectuer la moyenne des pixels situés dans une fenêtre de taille donnée et de l'affecter au pixel central, en réitérant cette opération en chaque point de l'image.

L'application d'un filtrage relève de la convolution d'une image par un noyau, assimilable au masque de la fonction de transfert du filtre, dont les coefficients peuvent être adaptés par l'utilisateur en fonction de l'objectif à atteindre.

Exemples de noyaux pour des filtres passe-bas :

$$\frac{1}{9} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \text{ ou filtre par la moyenne, } \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \text{ ou filtre gaussien} \quad [2.4]$$

Ce type de filtrage est plus adapté à l'élimination du bruit gaussien que du bruit impulsionnel. La mise en œuvre du filtre par la moyenne peut être optimisée en conservant en mémoire les traitements appliqués aux points du voisinage déjà traités ; pour le filtre gaussien, il se décompose en deux filtres monodimensionnels. Ces artifices permettent de réduire le nombre d'opérations à effectuer pour réaliser le filtrage.

L'utilisation de tels filtres entraîne une atténuation des fréquences spatiales élevées, d'autant plus forte que la taille de la fenêtre est grande. Il convient donc de prendre garde à ne pas altérer les objets que l'on souhaite conserver dans l'image, car leurs contours deviennent de moins en moins significatifs au fur et à mesure de ce filtrage. De plus, ces traitements génèrent des effets de bord d'autant plus larges que la taille de la fenêtre est importante ou que le nombre d'itérations appliquées est important. Ces filtres dégradent donc considérablement les contours et rendent l'image floue, à quelques exceptions près comme lorsque le signal est stationnaire et le bruit gaussien.



Figure 2.10. À droite, une image numérisée à 21 μm sur une photographie aérienne. Les zones homogènes, bruitées, contiennent une texture aléatoire qui ne constitue pas un signal utile. Au centre, l'image originale après le filtre gaussien. Si le bruit a été atténué, les contours sont également devenus flous. À gauche, une image d'une caméra numérique : les zones sont naturellement faiblement bruitées en raison de la nature du capteur et de sa très bonne dynamique.

2.1.2.4. Filtrage non linéaire dans le domaine spatial

Les filtrages linéaires ont pour inconvénient principal de traiter de la même façon les parties du signal porteuses d'information et de bruit, ce qui justifie parfois le recours à des traitements non linéaires, plus aptes à atténuer les valeurs aberrantes, dont la teinte est trop éloignée des pixels voisins. Également appelé filtrage de rang, ces méthodes remplacent le pixel central ou courant par l'une des valeurs triées des pixels du voisinage V_{xy} en (x, y) : pixel de valeur minimale, maximale ou médiane...

Ce type de méthodes est plus robuste par rapport au bruit d'image et présente l'avantage d'attribuer une valeur existante sans en calculer une nouvelle.

2.1.2.5. Filtre médian

Dans ce traitement, la valeur attribuée au pixel est la médiane statistique des valeurs contenues dans la fenêtre de filtrage, dont le nombre d'éléments doit donc être impair. La médiane est en effet le meilleur estimateur de la valeur moyenne d'un ensemble d'échantillons, moins sensible aux valeurs aberrantes, dont il est ici supposé qu'elles correspondent aux valeurs extrêmes de l'histogramme à l'intérieur du masque de filtrage.

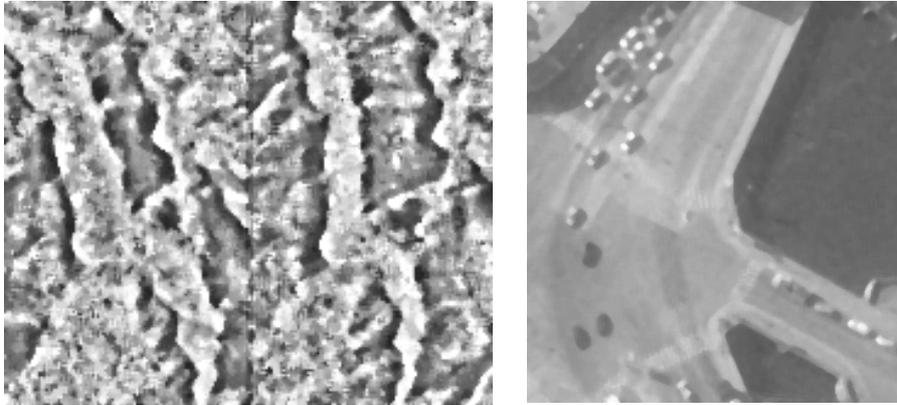


Figure 2.11. *Utilisation du filtre médian sur les images brutes utilisées précédemment. Il présente comme principal défaut d'arrondir les coins et de lisser les lignes de largeur inférieure à la moitié de celle du filtre*

Ce traitement permet d'éliminer les valeurs isolées qui correspondent à des pics de bruit impulsionnel et préserve plus efficacement les contours des objets que les filtres linéaires. Néanmoins, les déformations induites par ce filtre sur les objets s'accroissent avec la taille de la fenêtre utilisée. Ce filtrage est donc surtout recommandé lorsque les images présentent un bruit de type impulsionnel, donc à des variations importantes et de faible étendue. Certaines implémentations utilisent un masque en forme de croix centrée sur le pixel courant, de préférence à la forme carrée pour limiter les déformations dans l'image.

2.1.2.6. Filtrage adaptatif

Des filtres sont dits adaptatifs dans les cas suivants :

- les coefficients pondèrent chaque terme du filtre par la moyenne, par un terme qui décroît avec la similarité entre le pixel considéré et le pixel central de la fenêtre ;
- la fenêtre est choisie en forme ou en taille.

Le filtre de Nagao, relativement peu connu, donne pourtant de très bons résultats dans le cas d'images aériennes, en particulier en milieu urbain. Il se structure en trois étapes :

- une fenêtre 5 x 5 entourant le pixel central est successivement découpée en 9 fenêtres de 9 pixels chacune (*cf.* figure 2.12);

Les deux premiers modèles sont chacun déclinés en quatre schémas déduit de celui présenté par rotations successives de 90° , 180° et 270° .

- l'homogénéité de chaque fenêtre est mesurée à l'aide d'un indicateur du type variance radiométrique ;
- le pixel central est remplacé par la moyenne de la plus homogène des neuf zones.

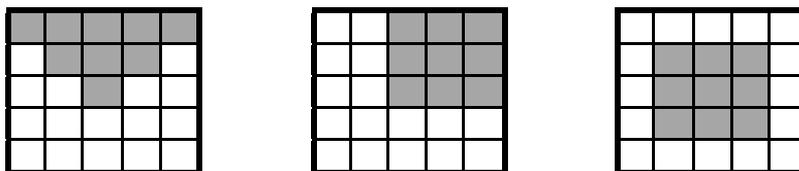


Figure 2.12. *Filtre de Nagao*

Le filtre obtenu lisse fortement les textures sans trop altérer les contours des objets dans les images.

Le filtre de Nagao présente le défaut de lisser les traits les plus fins de l'image. Des variantes algorithmiques existent donc pour trouver des indicateurs plus pertinents de la mesure de l'homogénéité. La taille des fenêtres utilisée, le type de pondération retenu et la forme des masques d'application peuvent être adaptés de manière empirique de manière à obtenir le meilleur filtre pour la préservation des lignes, des coins ou la suppression du bruit.

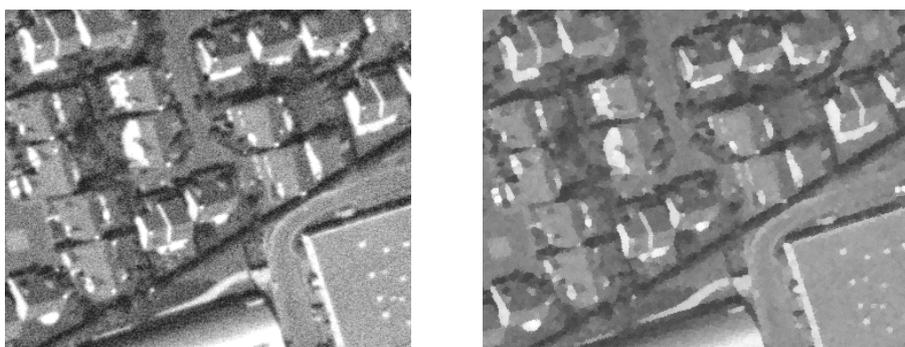


Figure 2.13. *Application d'un filtrage de Nagao*

2.1.2.7. Filtrage dans le domaine fréquentiel

Le filtre dans le domaine fréquentiel s'effectue en trois étapes :

- calcul de $F(u, v) = TF(I(x, y))$; en général, c'est la Transformée de Fourier rapide qui est utilisée sur l'image de base ;
- multiplication par la fonction de transfert $H(u, v)$ d'un filtre *ad hoc* : $G(u, v) = H(u, v) F(u, v)$;
- obtention de l'image filtrée par la Transformée de Fourier inverse de $G(u, v)$.

Le temps de calcul devient un paramètre prépondérant qu'il convient de prendre en compte avant toute utilisation dans un contexte opérationnel. La pratique a consacré parfois l'usage, en lieu et place des transformées de Fourier directe et inverse, d'un masque carré dont les coefficients sont les éléments discrétisés de la fonction de transfert du filtre du domaine fréquentiel. Les filtres de Butterworth, passe-bas exponentiel et gaussien sont les plus fréquemment évoqués, même si pour l'instant, relativement peu d'applications en photogrammétrie numérique les utilisent.

Le filtrage homomorphique est également utilisable pour corriger les irrégularités d'éclairage d'un objet. En partant du constat qu'une image $I(x, y)$ peut s'écrire comme le produit de $e(x, y) r(x, y)$, où e est la fonction d'éclairage qui génère plutôt des basses fréquences lors du calcul de la transformée de Fourier et r est la réflectance qui génère plutôt des hautes fréquences de par la nature des objets qui la composent.

Les deux effets sont rendus additifs en créant une image logarithmique telle que :

$$\text{Ln}(I(x, y)) = \text{Ln}(e(x, y)) + \text{Ln}(r(x, y)). \quad [2.5]$$

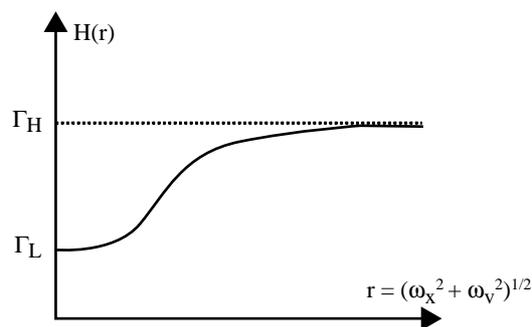


Figure 2.14. Filtrage homomorphique

Le filtrage homomorphique va rehausser les hautes fréquences et réduire les basses fréquences au point de réduire les variations de l'éclaircissement tandis que les détails seront rehaussés, permettant ainsi une meilleure observation dans les zones sombres de l'image.

La fonction de transfert est de la forme :

$$H(\omega_x, \omega_y) = \frac{1}{1 + \exp^{-s(\sqrt{\omega_x^2 + \omega_y^2} - \omega_0)}} + A \text{ avec des valeurs comme } s=1, \omega_0 = 128$$

et $A = 10$, ces paramètres étant reliés comme suit aux paramètres Γ_H et Γ_L par

$$\Gamma_L = \frac{1}{1 + \exp^{s\omega_0}} + A, \Gamma_H = 1 + A.$$

2.1.3. Rehaussement des contours

Les points de lignes de contours constituent la base majoritaire de l'information contenue dans les images, probablement plus que les textures qui jouent pourtant un rôle essentiel, en particulier pour la perception du relief. La reconnaissance d'un objet se fait le plus souvent uniquement à partir de ses contours. La difficulté réside dans la notion subjective que l'utilisateur présume pour définir l'utilité et l'importance d'un contour. La multiplicité des détecteurs de contours résulte de la variété des images étudiées et du type d'application poursuivi.

2.1.3.1. Les méthodes du gradient

Cette première catégorie est basée sur la recherche de la dérivée première de l'intensité lumineuse. Les opérateurs rencontrés le plus souvent sont construits à partir des modèles suivants :

$$\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{pmatrix}$$

Gradient

Roberts

Prewitt

Sobel

Kirsh

[2.6]

Les trois derniers opérateurs mentionnés ci-avant présentent l'avantage d'être moins sensibles au bruit que la dérivée traditionnelle mais aussi, malheureusement, d'être moins précis. Comme ils sont chacun directionnels, il est habituel de construire les filtres similaires pour détecter les contours dans les huit directions possibles.

La convolution par des opérateurs de Sobel :

$$\begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \begin{pmatrix} -1 & -2 & -1 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{pmatrix} \dots$$

(qui seraient au nombre de 8 en déplaçant l'ensemble des coefficients d'un quart de tour pour chaque nouveau filtre) permet de détecter les zones de contours dans une direction donnée.

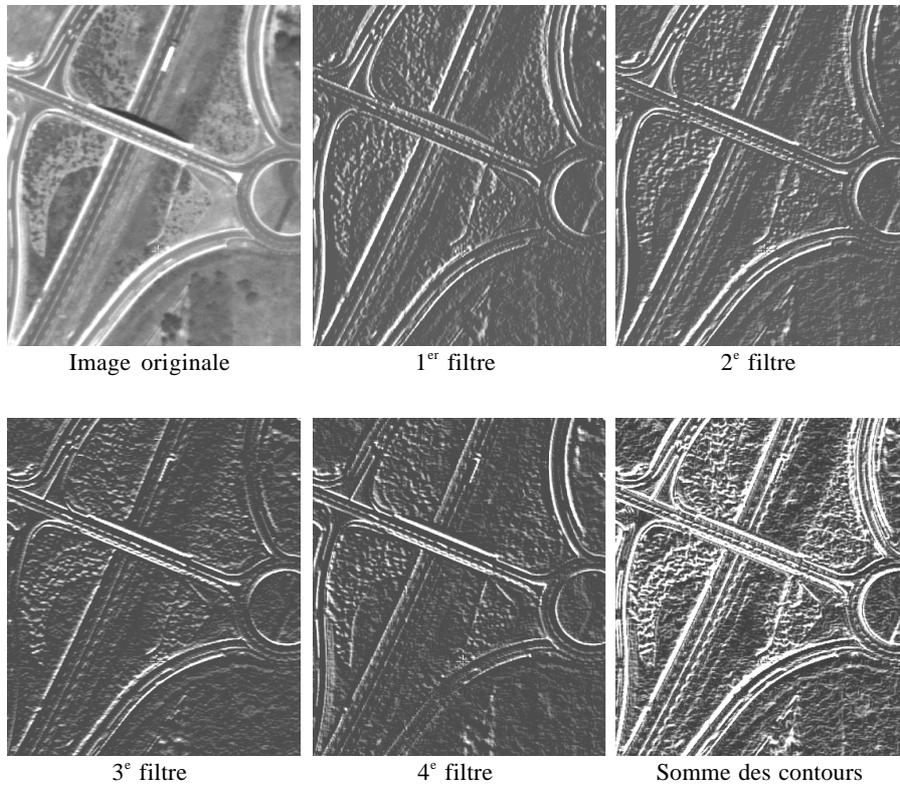


Figure 2.15. Filtrage par opérateur de Sobel

La superposition des quatre images obtenues avec les filtres proposés (cf. figure 2.15) permet d'obtenir une image de l'amplitude du gradient ; pour chaque pixel $I(x, y)$, le maximum trouvé dans les images intermédiaires aux coordonnées (x, y) est retenu. La présence d'un contour peut être décidée par rapport à un seuil sur cette valeur.

On aura souvent intérêt à réduire le bruit d'image avec des méthodes comme le filtre médian ou le filtre de Nagao, ce qui n'empêchera pas les contours d'être de qualité très moyenne car bruités, épais et non continus.

2.1.3.2. Les méthodes de passage par zéro de la dérivée seconde

Cette seconde catégorie est basée sur l'étude du passage par zéro des dérivées secondes de l'intensité lumineuse afin de détecter les contours. Si les zéros de la dérivée seconde constituent un réseau de lignes fermées, les dérivées secondes sont généralement très bruitées. Pour pallier cet inconvénient, il est possible d'élargir le support du masque utilisé, de ne sélectionner que les passages par zéro s'accompagnant d'une forte pente du signal ou bien encore de filtrer les contours obtenus après la détection.

Il existe plusieurs types de laplaciens :

$$\begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix} \quad \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix} \quad \begin{pmatrix} -1 & 2 & -1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix} \quad [2.7]$$

Après l'application de ces masques sur l'image, on détecte les points de contour par la détection des passages par zéro de l'image obtenue. L'unicité du passage par zéro fournit directement des contours d'un pixel d'épaisseur.

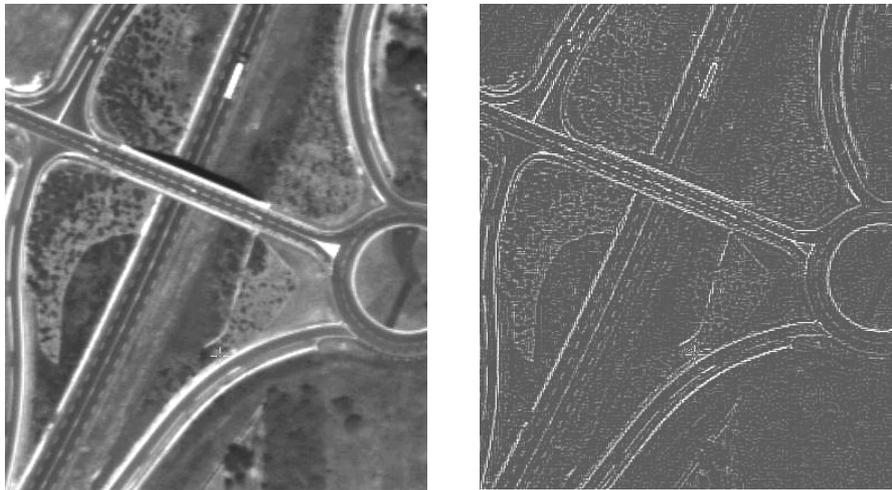


Figure 2.16. Filtrage laplacien

On montre que l'opérateur le plus approprié est constitué par le laplacien de gaussienne défini par :

$$G(x, y) = -\frac{1}{2\pi\sigma^4} \left(2 - \frac{x^2 + y^2}{\sigma^2} \right) \times e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad [2.8]$$

Une autre méthode calcule les passages par zéro de la dérivée seconde directionnelle en approximant l'intensité dans une fenêtre par un polynôme dont les coefficients sont calculés en convoluant l'image avec des masques. On peut ensuite calculer les dérivées secondes dans la direction du gradient et trouver les points de contour comme les passages par zéro de ces dérivées secondes directionnelles. Pour un point de coordonnées (i,j), les coefficients approximant l'intensité dans la base :

$$\left(1, i, j, (i^2 - \frac{2}{3}), ij, (j^2 - \frac{2}{3}), (i^2 - \frac{2}{3})j, i(j^2 - \frac{2}{3}), (i^2 - \frac{2}{3})(j^2 - \frac{2}{3}) \right)$$

sont les suivants :

$$\begin{aligned} & \frac{1}{9} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \frac{1}{6} \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \frac{1}{6} \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}, \frac{1}{6} \begin{pmatrix} -1 & -1 & -1 \\ -2 & -2 & -2 \\ 1 & 1 & 1 \end{pmatrix}, \frac{1}{4} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \\ & \frac{1}{6} \begin{pmatrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \end{pmatrix}, \frac{1}{4} \begin{pmatrix} -1 & 0 & 1 \\ 2 & 0 & -2 \\ -1 & 0 & 1 \end{pmatrix}, \frac{1}{4} \begin{pmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ 1 & -2 & 1 \end{pmatrix}, \frac{1}{4} \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix} \quad [2.9] \end{aligned}$$

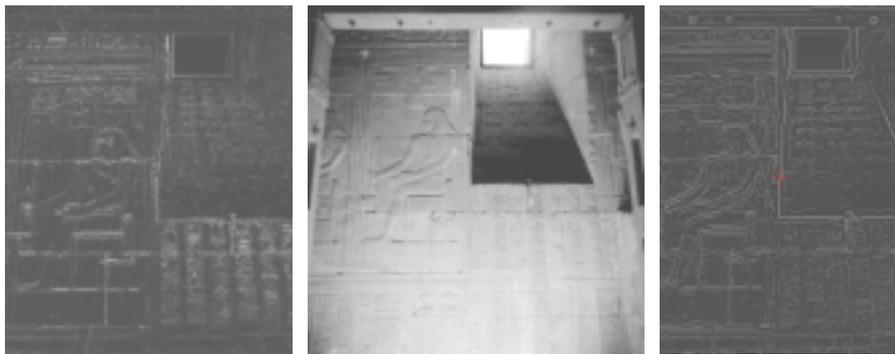


Figure 2.17. De gauche à droite : image des contours obtenue par un Laplacien, image originale et contours obtenus par le gradient de Deriche

Le coût de calcul de ces derniers traitements est suffisamment important pour que des approches de compromis entre rapidité et performances aient été développées ; il en va ainsi du filtre de Deriche [DER 87], qui possède une réponse impulsionnelle infinie de la forme $f(x) = -c \times \exp^{-\alpha|x|}$ permettant donc une implémentation au moyen de filtres récursifs séparables. Le filtrage 2D est obtenu par l'action de deux filtres croisés en x et en y et le filtrage du bruit à l'aide d'une fonction de prolongement (généralement une gaussienne de même écart-type que celle considérée pour le détecteur).

2.1.4. Conclusion

L'amélioration d'image peut souvent être traitée par des techniques de convolution avec des filtres appropriés, d'extension finie, qui permettent ainsi une implémentation sur l'image comme des opérateurs linéaires. L'amélioration d'une image cherche à la rendre esthétique suivant des critères subjectifs.

L'amélioration d'image peut utiliser soit :

- des techniques de manipulation des niveaux de teintes en utilisant les propriétés de l'histogramme auquel l'utilisateur peut donner la forme qu'il désire ;
- des techniques de filtrage en vue de réduire le bruit et de rehausser les contours dans l'image. Les méthodes les plus courantes sont souvent non linéaires.

2.2. Compression des images numériques²

2.2.1. Intérêt de la compression des images numériques

La numérisation des images offre un grand nombre d'avantages :

- traitement possible par logiciel d'où une grande souplesse ;
- grande fiabilité du stockage (sur CD-ROM, disques durs...) ;
- transmission sans erreur (grâce aux codes correcteurs d'erreur).

Néanmoins, elle présente l'inconvénient de générer le plus souvent de gros volumes de données. À titre d'exemple, une image de télédétection spatiale acquise par le satellite SPOT dans la bande panchromatique représente un volume de : 6 000 lignes x 6 000 colonnes x 8 bits/pixel = 288 Mbits. Une photo aérienne classique, scannée à 14 μm , représente 2 048 Mbits.

2. Cette section a été rédigée par Gilles MOURY.

Bien entendu le premier objectif d'un ingénieur, lorsqu'il conçoit le système de transfert de données, devrait être de définir le nombre effectif de bits par pixel en tenant soigneusement compte du bruit de l'image, par exemple pour que le niveau de bruit attendu soit au niveau du dernier bit. Mais ce nombre est presque toujours pris plus important, pour être capable de traiter des situations significativement différentes de la situation nominale, ce qui signifie bien souvent un ou deux bits au-delà de l'optimum.

Compte tenu des limitations qui s'appliquent dans la plupart des systèmes sur les capacités de stockage et/ou de transmission, il est primordial de réduire au maximum la quantité de bits par pixel nécessaire pour représenter l'image. Comprimer, c'est choisir la représentation qui conduit au minimum de bits, tout en préservant dans l'image les informations nécessaires à l'utilisateur. L'efficacité de la compression sera mesurée par le taux de compression qui est le rapport entre le nombre de bits de l'image source sur le nombre de bits de l'image comprimée.

La compression des images est possible pour les raisons suivantes :

- les données images acquises par l'instrument de prise de vue présentent une redondance naturelle qui ne contribue pas à l'information et qu'il est possible d'éliminer avant stockage et transmission. On peut donc compresser efficacement les images sans aucune perte d'information (*compression dite réversible*). Néanmoins, nous verrons par la suite que ce type de compression est limité à des taux de compression relativement faibles (typiquement entre 1,5 et 2,5 suivant le contenu de l'image) ;

- de plus, l'utilisateur final des images ne s'intéresse, en général, qu'à une partie des informations véhiculées par l'image. C'est ce que l'on appellera l'information pertinente. Il sera donc possible de compresser plus efficacement encore les images en enlevant l'information non pertinente (*compression dite avec pertes*) et ceci à satisfaction égale de l'utilisateur. Les taux de compression pourront alors être beaucoup plus élevés (jusqu'à 50 dans certaines applications très peu exigeantes en terme de qualité). Bien entendu, le niveau de dégradation dû à la compression sera fonction du taux et de la performance de l'algorithme utilisé.

Pour chaque application, il y aura donc un compromis à trouver, en fonction des limitations des ressources (stockage, transmission), entre la satisfaction des utilisateurs (la qualité des images étant inversement proportionnelle au taux de compression utilisé) et la quantité d'images qui pourront être stockées et/ou transmises.

2.2.2. Critères de choix d'un algorithme de compression

Les critères à prendre en compte pour le choix d'un algorithme de compression et d'un taux de compression sont très divers et dépendent en partie de l'application visée. Parmi les plus génériques, on peut citer :

- le type d'images : images de type photographique comportant un grand nombre de niveaux de gris ou de couleurs (8 à 24 bits par pixel), ou bien images artificielles comportant quelques niveaux de gris seulement (exemple : images binaires de type fax). Ces deux types d'images requièrent l'utilisation d'algorithmes très différents. Pour ce qui est des algorithmes normalisés, on peut citer : le standard JPEG pour les images photographiques, le standard JBIG pour les images artificielles à peu de niveaux de gris et le standard T4 et/ou T6 pour la télécopie noir et blanc. Nous nous intéresserons par la suite uniquement aux images de type photographique ;

- le niveau de qualité requis par l'utilisateur : peut aller d'un besoin de compression strictement réversible (cas de certaines applications en imagerie médicale) à des besoins en qualité très faibles (cas de certaines applications de transmission d'images sur Internet). Le niveau et la nature des pertes tolérables seront bien entendu à affiner avec l'utilisateur par exemple au travers de campagnes d'évaluation pour lesquelles plusieurs niveaux et types de dégradations seront simulés sur des images représentatives de l'application visée. Le type de dégradation (artefacts) va être plutôt fonction de l'algorithme et le niveau de dégradation plutôt fonction du taux de compression ;

- le type d'algorithme : normalisé ou propriétaire. Les principaux avantages des algorithmes normalisés sont bien entendu la compatibilité et la pérennité avec la garantie de compatibilité ascendante. L'inconvénient majeur est l'évolution lente des normes qui fait qu'une norme finalisée est très rarement la solution la plus performante disponible ;

- le type de transmission ou d'accès à l'image décompressée : séquentielle ou progressive. Dans le premier cas, l'image est transmise en bloc à la résolution maximale ; dans l'autre cas, on transmet d'abord une version basse résolution (donc très compacte) qui permet à l'utilisateur, par exemple, de choisir dans une base de données, puis après sélection une version pleine résolution. Les algorithmes par transformée (DCT, ondelettes...) permettent, entre autres, les transmissions progressives.

2.2.3. Quelques éléments théoriques

2.2.3.1. Corrélation des données image

Dans une image naturelle (par opposition à une mire artificielle), la corrélation entre pixels voisins est très élevée et décroît en fonction de la

distance euclidienne entre les pixels. La corrélation diminue très vite avec la distance et devient souvent négligeable à 5 - 10 pixels de distance. Pour comprimer efficacement une image, et donc éviter de coder et transmettre plusieurs fois la même information, la première opération à réaliser est toujours de décorrélérer localement l'image. Nous verrons plusieurs types de décorrélateur par la suite. La décorrélation la plus simple à réaliser consiste, au lieu de coder chaque pixel $p(i,j)$ indépendamment de ses voisins, à coder la différence: $p(i,j) - p(i,j-1)$ avec i indice de ligne et j indice de colonne.

2.2.3.2. Notion d'entropie

L'entropie est une mesure de la quantité d'informations contenues dans un jeu de données. Elle est définie de la façon suivante :

- considérons une image codée sur k bits/pixel. Chaque pixel peut prendre 2^k valeurs entre 0 et 2^k-1 ;
- l'entropie d'ordre 0 de l'image (notée $H_0(S)$) est donnée par la formule :

$$H_0(S) = \sum_{i=0}^{i=2^k-1} P_i \cdot \log_2 \frac{1}{P_i} \quad [2.10]$$

où P_i est la probabilité pour qu'un pixel de l'image prenne la valeur i .

La théorie de l'information [RAB 91] montre que $H_0(S)$ donne, pour une source sans mémoire donc décorrélée, le nombre moyen minimal de bits par pixel avec lequel il est possible de coder l'image sans perdre d'information. En d'autres termes, le taux de compression maximal qu'il sera possible d'atteindre en compression réversible est donné par : $CR_{\max} = \frac{k}{H_0(S)}$.

Pour illustrer la notion d'entropie, prenons quatre types d'images différents :

- une image uniforme où tous les pixels ont la même valeur, $H_0(S)=0$, cette image ne contient aucune information ;
- une image binaire noire-blanche du type de celles traitées par les télécopieurs :

$$H(S) = P_{\text{blanc}} \cdot \log_2 \frac{1}{P_{\text{blanc}}} + P_{\text{noir}} \cdot \log_2 \frac{1}{P_{\text{noir}}} \quad [2.11]$$

on a donc $H(S) < 1$. En pratique, $P_{\text{noir}} \ll P_{\text{blanc}}$ donc $H(S) \ll 1$, ce qui explique que les algorithmes de compression réversibles utilisés dans les télécopieurs ont des taux de compression moyens supérieurs à 100 ;

- l'image SPOT de Gênes sur 8 bits en figure 2.22 : $H_0(S) = 6,97$;
- une image (8 bits) de bruit blanc saturé dans laquelle toutes les valeurs sont équiprobables (histogramme plat), $H_0(S) = 8$. Il n'est donc pas possible de compresser cette image de façon réversible.

2.2.3.3. Compression réversible

La compression réversible est limitée par le taux de compression $CR_{\max} = k/H_0(S)$. Pour réaliser une compression réversible, on va d'abord décorréler l'image de façon à réduire l'entropie $H_0(S)$. On peut illustrer l'effet de la décorrélation sur l'entropie, sur l'image de la figure 2.22 :

- entropie de l'image source : $H_0(S) = 6,97$;
- entropie de l'image après décorrélation de type :
[p (i, j) – p (i, j-1)] : $H_0(S) = 5,74$.

Plus la décorrélation sera efficace, plus l'entropie de l'image après décorrélation sera faible et plus le taux de compression atteignable par codage réversible des données décorréelées sera élevé. Néanmoins, les taux de compression atteignables restent relativement faibles (de l'ordre de 1,5 à 2,5 sur des images de télédétection du type de celles prises par le satellite SPOT).

2.2.3.4. Compression avec pertes

Lorsque l'on recherche des taux de compression supérieurs à CR_{\max} , ce qui est le cas le plus courant, on est obligé d'introduire des pertes d'informations dans la chaîne de compression. Ces pertes d'informations sont en général réalisées par une quantification des données décorréelées. Pour une image et un algorithme de compression donnés, il existe une relation entre le taux de compression obtenu et l'erreur introduite dans l'image par la compression/décompression. Cette courbe, appelée « taux-dégradation » a l'allure typique donnée en figure 2.18.

La dégradation est souvent mesurée quantitativement par l'écart-type de l'erreur de compression, donné par la formule :

$$\sigma = \sqrt{\frac{1}{N.M} \sum_{i=1}^N \sum_{j=1}^M (p(i,j) - p'(i,j))^2} \quad [2.12]$$

pour une image N lignes, M colonnes et une erreur moyenne de compression nulle, ce qui est le cas le plus général ($p'(i, j)$ étant la valeur du pixel reconstruite après compression/décompression).

Le niveau maximal de dégradation tolérable par l'utilisateur des images fixe le taux de compression maximal utilisable sur la mission. Pour comparer les performances de deux algorithmes de compression, il importe de tracer les courbes « taux-dégradation » des deux algorithmes et ceci sur la (ou les) même(s) image(s), la performance d'un algorithme variant notablement d'une image à l'autre en fonction de l'entropie des données à compresser.

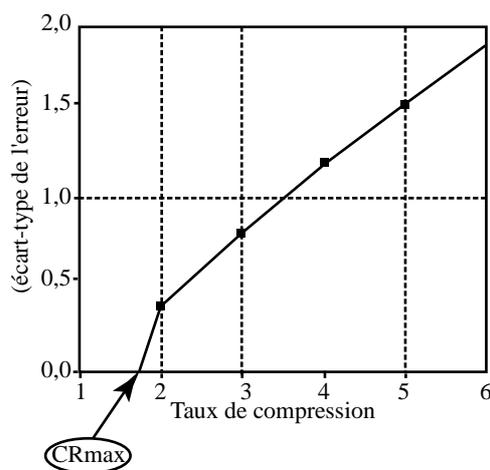


Figure 2.18. Courbe « taux-dégradation » typique

2.2.4. Présentation des divers types d'algorithmes

Une présentation complète des divers types d'algorithmes de compression peut être trouvée dans les références [RAB 91], [GUI 95], [RAO 96]. Nous nous bornerons ci-après à donner les principes généraux et à décrire quelques algorithmes parmi les plus utilisés.

2.2.4.1. Architecture générale

Tout système de compression peut se décomposer en trois modules distincts (cf. figure 2.19) : à savoir : la *décorrélation* de l'image source, la *quantification* des valeurs décorrélées et l'*affectation de codes* binaires.

Le décorrélateur permet de réduire la redondance contenue dans les données. En pratique, il existe un grand nombre de méthodes, plus ou moins locales, pour décorrélérer les pixels incidents. Nous en donnons quelques exemples par la suite (DPCM, DCT et transformée en ondelettes). Cette phase de traitement d'image est parfaitement réversible. À l'issue de cette première phase, le processus de

compression n'a donc pas encore effectivement commencé, mais les données ont été décorrélées pour que les traitements ultérieurs (quantification, codage) soient optimaux.

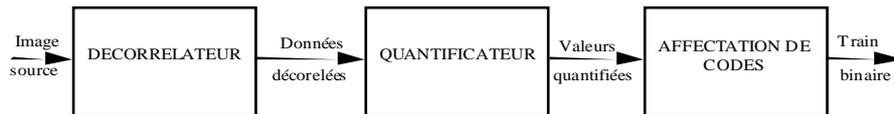


Figure 2.19. Schéma général d'un système de compression

Le quantificateur est l'organe essentiel du système de compression. En effet, c'est lui qui va diminuer effectivement la quantité d'information transmise en éliminant toute l'information non pertinente (vis-à-vis de l'utilisation qui est faite des images) contenue dans les données issues du décorrélateur. Rappelons que l'information non pertinente est une quantité qui dépend uniquement de (ou des) application(s) envisagée(s) : exemple, dans le cas des missions SPOT, la précision radiométrique n'est pas pertinente dans les zones de fortes transitions (contours très contrastés). La quantité d'information non pertinente éliminée par le quantificateur doit pouvoir varier en fonction de l'application. Nous voyons donc que le quantificateur joue le rôle d'organe de commande du système de compression en le rendant utilisable à différents débits. La quantification est la seule opération non réversible de la chaîne de compression.

L'affectation de codes (ou plus simplement codage) a pour rôle de produire un train binaire, représentatif des valeurs quantifiées, qui sera effectivement transmis ou stocké pour transmission ultérieure. Le taux de compression réellement atteint par le système ne peut être évalué de façon réaliste qu'à l'issue de ce module. Son rôle est d'affecter à chaque valeur quantifiée ou événement un code binaire qui pourra être déchiffré sans ambiguïté par le décodeur au sol. Cette affectation peut se faire de façon plus ou moins économique en terme de nombre de bits transmis. Les codes les plus efficaces sont les codes à longueur variable dont le principe est simple : affecter les codes les plus courts aux valeurs (ou événements) les plus probables. Un exemple de code à longueur variable est le code de Huffman, utilisé dans les standards JPEG et MPEG.

La mise en œuvre des diverses méthodes de compression ne fera pas toujours apparaître les trois modules précités. Certains modules pourront être regroupés en un seul (c'est le cas de la quantification et du codage pour les algorithmes utilisant le codage vectoriel). Le module de quantification peut également disparaître comme dans le cas d'une méthode de compression réversible.

2.2.4.2. Algorithmes réversibles

Il existe deux grandes classes d'algorithmes pour réaliser une compression réversible d'images. Les algorithmes universels capables de comprimer n'importe quel type de données et les algorithmes spécifiquement optimisés pour les images, ces derniers donnant des performances 20 à 30 % meilleures en termes de taux de compression.

2.2.4.2.1. Algorithmes universels

Le plus répandu est l'algorithme Lempel-Ziv (LZ, LZ77, LZW) utilisé dans les utilitaires zip, gzip, pkzip et dans les formats d'image tif, gif, png. Son principe consiste à repérer des séquences de symboles (caractères d'un texte, valeurs de pixel) qui se répètent dans le fichier. Ces séquences sont alors stockées dans un dictionnaire qui est remis à jour dynamiquement. Chaque séquence (de longueur variable) est codée par un code de longueur fixe dépendant de la taille du dictionnaire (exemple : 12 bits pour un dictionnaire de 4 096 éléments). Ce type d'algorithme tire parti de la corrélation de la source et s'adapte en temps réel à la statistique locale (par mise à jour du dictionnaire). Néanmoins, le codage des pixels étant linéaire suivant le balayage ligne, on ne tire pas parti des corrélations verticales dans l'image d'où une moins bonne performance que les algorithmes spécifiques qui tirent parti des corrélations dans les deux dimensions.

2.2.4.2.2. Algorithmes spécifiques

L'algorithme le plus performant est la nouvelle norme JPEG-LS [4]. Le schéma-bloc de cet algorithme est donné en figure 2.20. Il utilise un décorrélateur de type prédictif à deux dimensions (le plus souvent utilisé dans les algorithmes sans perte). La valeur du pixel courant x est prédite à partir d'une combinaison linéaire des pixels a, b, c (cf. figure 2.20) encodés précédemment. L'erreur de prédiction est ensuite codée à l'aide d'un code de Huffman adaptatif en fonction du contexte (a, b, c, d) du pixel courant (analysé par le module de modélisation du contexte). Un codage par plage est utilisé dans les zones complètement uniformes.

Sur l'image SPOT de Gênes (cf. figure 2.22), on obtient la comparaison suivante entre algorithmes universels et spécifiques, soit un gain d'environ 30 % sur le taux.

Algorithme	Taux de compression
JPEG-LS	1,65
Lempel-Ziv (gzip -9)	1,27

Tableau 2.1. Taux de compression sur l'image 2.22 pour différents algorithmes

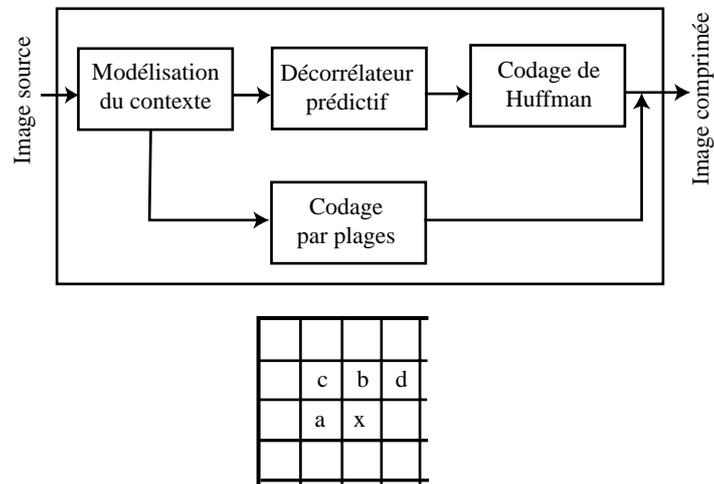


Figure 2.20. Schéma-bloc de l'algorithme JPEG-LS & contexte pour la modélisation et la prédiction

2.2.4.3. Algorithmes avec pertes

Ces algorithmes sont classés, du plus simple au plus complexe, suivant les trois principaux types de décorrélateurs utilisés : différentiel (DPCM), transformée cosinus (DCT), transformée en ondelettes (Multirésolution).

2.2.4.3.1. DPCM

Le décorrélateur différentiel DPCM consiste à prédire la valeur du pixel à coder à partir d'une combinaison linéaire des valeurs de ses voisins déjà codés (cf. figure 2.20). L'erreur de prédiction est ensuite quantifiée à l'aide d'une loi non uniforme adaptée à la statistique Laplacienne de ce résidu de prédiction. Ce type d'algorithme, utilisé sur SPOT 1/2/3/4 à un taux de 1.33, présente l'intérêt d'être très simple et de concentrer les erreurs de compression dans les zones de fortes transitions radiométriques (contours bien contrastés), zones dans lesquelles la valeur absolue de la radiométrie importe peu dans les applications de télédétection. Son désavantage est sa mauvaise performance en terme de « taux-dégradation » pour des taux moyens à élevés (> 5) (cf. figure 2.26) due à une décorrélation trop locale du signal. Pour la compression avec pertes, le décorrélateur DPCM a donc été abandonné au profit de la DCT.

2.2.4.3.2. DCT

La DCT (*Discrete Cosine Transform*) est le décorrélateur le plus couramment utilisé en compression d'images. Elle est à la base de nombreux standards de compression dont les plus connus sont :

- ISO/ JPEG pour les images fixes [International Standard ISO-10918] ;
- CCITT H261 pour la vidéoconférence ;
- ISO/MPEG pour la compression vidéo.

La DCT est une des transformées unitaires (cf.[RAB 91] pour plus de précision sur la définition et l'intérêt des transformées unitaires en compression) utilisées en compression pour réaliser la décorrélation. Cette transformation opère sur des blocs de $n \times n$ pixels ($n = 8$ en général) et réalise sur ces blocs une transformée de Fourier de période $2n$ dont on ne conserve que la partie réelle. Le schéma général d'une compression à base de transformée unitaire DCT est donné en figure 2.21.

La transformation DCT directe d'un bloc 8×8 est donnée par la formule suivante, où les $p(i,j)$ sont les pixels de l'image source et les $F(u, v)$ sont les coefficients DCT représentatifs du contenu spectral du bloc aux différentes fréquences spatiales :

$$F(u, v) = \frac{1}{2} C_u C_v \sum_{i=0}^7 \sum_{j=0}^7 p(i, j) \cdot \cos \frac{(2i+1)u\pi}{16} \cdot \cos \frac{(2j+1)v\pi}{16}$$

avec : C_u (resp. C_v) = $\frac{1}{\sqrt{2}}$ pour $u = v = 0$ et $C_u = C_v = 1$ sinon [2.13]

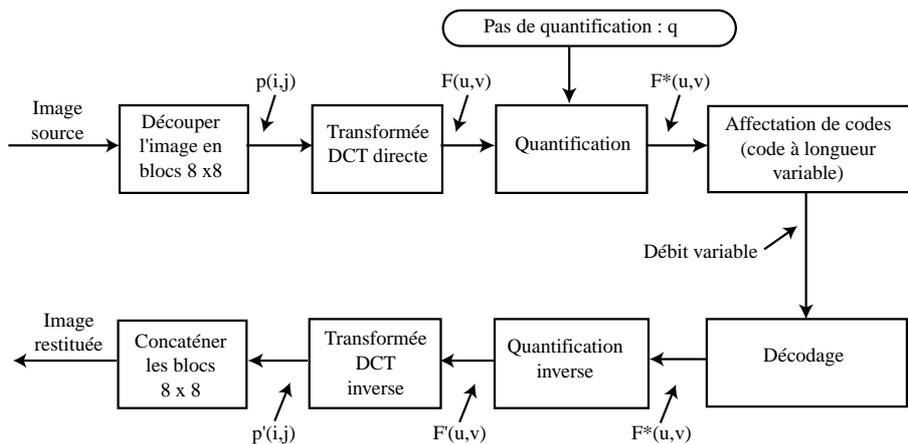


Figure 2.21. Chaîne de compression par transformée DCT

Dans le domaine transformé (domaine des fréquences spatiales) :

- les coefficients $F(u,v)$ obtenus sont bien décorrélés contrairement aux pixels sources du bloc ;
- les coefficients qui ont une amplitude non négligeable se trouvent statistiquement concentrés dans une région restreinte du plan transformé, ce qui facilite grandement le codage ultérieur après quantification de ces coefficients.



Figure 2.22. Extrait d'image SPOT panchromatique sur la ville de Gênes – original non comprimé

Figure 2.23. Image de Gênes comprimée avec un algorithme type JPEG à un taux de 8 (rmse = 6,8)

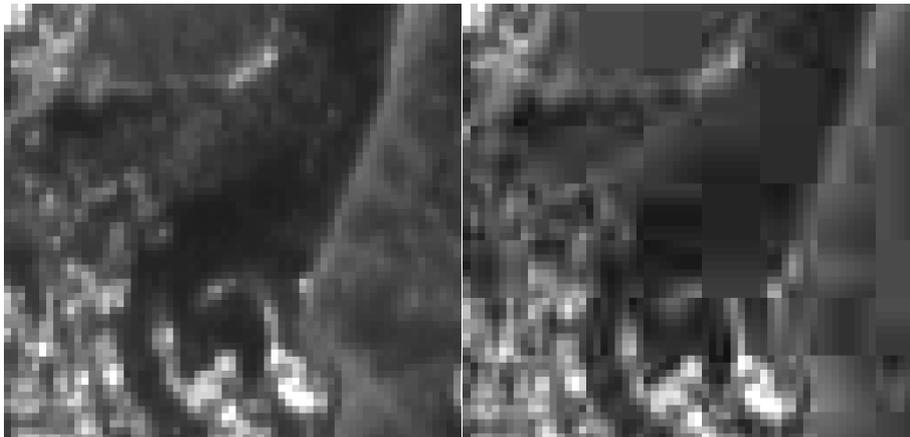


Figure 2.24. Zoom sur l'image de Gênes non comprimée

Figure 2.25. Zoom sur l'image de Gênes comprimée JPEG à un taux de 8 (effet de blocs 8 x 8 bien visible)

La quantification est une simple division par un pas de quantification q . Plus q sera grand, plus le taux de compression sera élevé et inversement. Le taux et donc l'erreur sur les données reconstruites sont donc contrôlés par le choix du pas. En sortie de quantification, un grand nombre de coefficients ($F^*(u, v)$) sont nuls. On utilise donc un codage de Huffman sur des événements de type plage-longueur (*run-length*) adapté à cette statistique. Ce codage, particulièrement efficace, a été optimisé dans le cadre des travaux sur le standard ISO/JPEG. Ce standard, dont une description détaillée est donnée en [PEN 93], est très largement utilisé.

Le principal défaut des algorithmes à base de DCT est lié au fait que chaque bloc 8x8 est codé indépendamment de ses voisins, ce qui crée des problèmes de raccordement de blocs après décompression. Ce phénomène (appelé effet de blocs), invisible à l'œil pour des taux de compression faibles (< 8), peut être gênant pour certains traitements informatiques de l'image. Il est illustré dans les figures 2.22 à 2.25 qui présentent respectivement : l'image source, l'image comprimée DCT/JPEG à un taux de 8, un zoom sur l'image source et le même zoom sur l'image comprimée faisant apparaître l'effet de blocs. Pour faire disparaître cet effet de blocs et améliorer encore la décorrélation par rapport à la DCT, on a recours à des algorithmes basés sur la décomposition en sous-bandes (par transformée en ondelettes) décrits ci-après.

2.2.4.3.3. Transformée en ondelettes

L'approche des techniques de codage en sous-bandes est identique à celle du codage par transformée unitaire par bloc (DCT ou autres) : décomposer le signal en composantes de différentes fréquences afin de coder chacune d'elles séparément. Les moyens mis en œuvre pour obtenir cette décomposition sont cependant différents : dans la méthode par bloc une matrice 8 x 8 de transformation fournit les raies de fréquences différentes, alors qu'un véritable filtrage de l'image entière doit être réalisé dans la technique sous-bandes. Cette décomposition en N sous-bandes s'effectue au moyen d'un filtrage hiérarchique de l'image. Les filtres utilisés sont déterminés à partir de la théorie des ondelettes [VER 95]. On réalise ainsi une transformée en ondelettes de l'image.

Les avantages des algorithmes de décomposition en sous-bandes par ondelettes sont :

- la décomposition de l'image en différentes fréquences spatiales est faite globalement sur l'image (par une fenêtre mobile de filtrage) et non par bloc 8 x 8 comme dans la DCT. De ce fait, il n'y a pas les problèmes de raccordement aux frontières de blocs sur l'image décomprimée ;

– les performances « taux-dégradation » sont bien meilleures pour des taux supérieurs à 8 (typ.) : amélioration de 30 % à 50 % du taux de compression à iso-dégradation (cf. figure 2.26) ;

– la transmission progressive par résolution et le débruitage sont très facilement intégrables dans ce type de schéma de compression.

Pour ces différentes raisons, le comité de normalisation JPEG a été réactivé pour définir la norme JPEG2000 basée sur la transformée en ondelettes. Cette future norme ISO devrait être finalisée en 2000 (<http://www.jpeg.org>) et intègre un grand nombre de fonctions (accès aléatoire au fichier comprimé, transmission progressive, robustesse aux erreurs de transmission, compatibilité ascendante JPEG...).

Pour ces algorithmes à base de transformée en ondelettes, les techniques de quantification/codage les plus performantes sont de type quantification par plan de bits et codage par arbre de zéros. Les algorithmes représentatifs de l'état de l'art sont EZW [EMB 93] et SPIHT [IEE 96].

2.2.4.3.4. Comparaison

On compare les performances « taux-dégradation » des trois types de décorrélateurs (DPCM, DCT, ondelettes) au travers de trois algorithmes représentatifs, sur l'image SPOT de Gênes (cf. figure 2.22) :

- DPCM : algorithme JPEG-LS en mode quasi sans perte [International Standard ISO-14495-1 (JPEG-LS)] ;
- DCT : algorithme JPEG mode « baseline » [International Standard ISO-10918 (JPEG)] ;
- ondelettes : algorithme SPIHT [IEE 96].

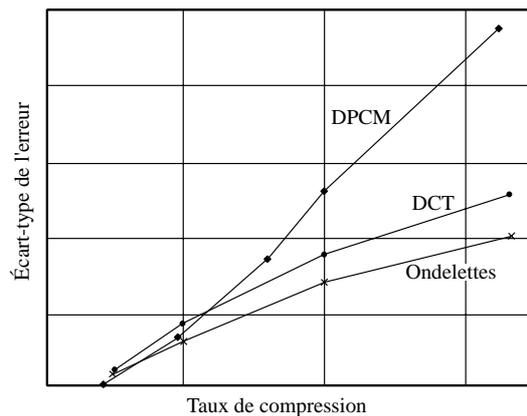


Figure 2.26. Courbes « taux-dégradation » comparatives DPCM, DCT, ondelettes

2.2.5. Compression multispectrale

Les algorithmes présentés jusqu'à maintenant ne tirent parti que de la corrélation intra-image (ou corrélation spatiale). Dans le cas d'une image multispectrale (ou couleur), il est possible d'exploiter également la corrélation entre les différentes bandes spectrales. Dans le cas des images SPOT par exemple, qui comportent 4 bandes (b1, b2, b3, mir), les bandes b1, b2 sont très fortement corrélées sur certains types de paysages dont la réponse spectrale varie peu.

Il est donc intéressant de réaliser successivement une décorrélation spectrale puis une décorrélation spatiale de l'image (l'ordre inverse est également possible). Une analyse complète des différentes techniques utilisables peut être trouvée dans [IEE 95]. Le décorrélateur spectral peut être du type prédictif ou par transformée orthogonale. La transformée optimale pour réaliser cette opération est la transformée de Karhunen-Loève (KLT). La transformée permet de passer d'un repère spectral $[b_i]$ où les composantes sont fortement corrélées, à un repère $[d_i]$ où les composantes sont décorréliées et l'information (énergie) de l'image concentrée sur peu de composantes. Un exemple d'une telle transformation est celui utilisé en télévision numérique (standard CCIR 601 [ITU 82]) : passage du repère $[RV,B]$ au repère $[Y,C_b,C_r]$, 80 % de l'information est concentré sur le signal de luminance Y ; les signaux de différences de chrominance (C_b, C_r) portant peu d'information peuvent dès lors être sous-échantillonnés et/ou comprimés à des taux plus élevés que la luminance.

2.2.6. Perspectives

Au-delà de la compression par transformée en ondelettes maintenant largement utilisée et en passe d'être normalisée (JPEG2000), différentes solutions sont actuellement à l'étude pour tenter d'améliorer encore les performances « taux-dégradation » des algorithmes. Parmi celles-ci, on peut citer :

- la compression par fractales [BAR 93] qui, couplée à la décomposition en sous-bandes par ondelettes, peut permettre de progresser dans la gamme des taux forts (>30) ;

- la compression sélective qui consiste à détecter dans l'image les zones d'intérêt pour l'utilisateur, de façon à leur appliquer un taux de compression préservant la qualité de l'image, tandis que le reste de l'image peut être comprimé avec un taux beaucoup plus fort (exemple : sur les images SPOT, les zones d'intérêt sont les zones hors nuages). La compression sélective ouvre la voie à l'utilisation de taux beaucoup plus élevés dans nombre d'applications. Toute la difficulté réside dans la fiabilité du module de détection.

2.3. Utilisation du GPS en photogrammétrie

2.3.1. Utilisation du GPS pour les prises de vues aériennes³

2.3.1.1. Introduction

Dans le cadre de prises de vues aériennes, le système GPS (*Global Positioning System*) ou tout autre système équivalent (GLONASS, ou le futur GALILEO), peut intervenir dans trois phases distinctes, la navigation en temps réel, la détermination de la trajectoire de l'avion, plus précisément la détermination de la position du sommet de prise de vue, et enfin dans la détermination de l'attitude de l'avion. Si ces trois aspects font effectivement appel au GPS, ils ne nécessitent pas le même type de mesure. L'utilisation la plus originale concerne la mesure d'attitude que nous préciserons plus loin.

Nous rappellerons brièvement ici qu'il existe deux types d'observations GPS :

- les pseudo-distances qui sont des mesures de temps de propagation de la modulation pseudo-aléatoire du signal ;
- la mesure de phase, qui se fait sur les ondes porteuses, et qui consiste à mesurer la phase du signal stable généré par le récepteur verrouillé sur le signal du satellite.

Il existe deux types de positionnement : le positionnement absolu et le positionnement relatif.

La précision du positionnement varie selon le type d'observations et de positionnement de quelques mètres à une précision relative de l'ordre de $2 \text{ mm} \pm 10^{-6} D$ à $10^{-8} D$ en mesure de phase.

En fin de chapitre, on trouvera un résumé de la terminologie employée pour les différents modes d'utilisation du GPS.

2.3.1.2. La navigation

La navigation est la partie qui pose *a priori* le moins de problème. Il s'agit d'utiliser les pseudo-distances. Au moment de la prise de vue, il est nécessaire d'avoir la position, en temps quasi réel, des sommets de prises de vues afin de vérifier que l'axe suivi par l'avion est le bon et que le taux de recouvrement réel est suffisant. En positionnement absolu, avec une précision qui est meilleure que la dizaine de mètres (brouillage SA désactivé), les conditions sont remplies. La précision obtenue est bien supérieure aux 40 m avec lesquels les pilotes peuvent

3. Cette section a été rédigée par Thierry DUQUESNOY.

tenir un axe de vol. Néanmoins, si une meilleure précision est requise, le mode différentiel est nécessaire : plusieurs solutions sont envisageables pour obtenir les corrections différentielles en temps réel. On peut utiliser des réseaux de stations qui émettent des corrections sur différentes bandes de fréquences. On peut disposer de son propre récepteur GPS sur une station fixe connue qui envoie les corrections. Enfin il existe des services de corrections différentielles *via* des satellites géostationnaires. Ce dernier procédé semble le mieux adapté pour les applications aéroportées.

Notons qu'en navigation différentielle, la vitesse de l'avion est connue à 0.1 nœud près, ce qui est suffisant pour les applications photogrammétriques.

Pour le recouvrement, il peut être nécessaire de connaître en même temps l'attitude de l'avion. Nous aborderons ces besoins au paragraphe 2.3.1.4.

Depuis mai 2000, la suppression du brouillage sélectif a beaucoup amélioré le positionnement en point isolé et en temps réel (qui est ainsi passé de 100 m à moins de 10 m d'écart type), et dans ces conditions le recours à des corrections différentielles temps réel est devenu peu utile pour le suivi des plans de vol.

2.3.1.3. La détermination des sommets de prises de vues pour l'aérotriangulation

Nous avons vu que les mesures de pseudo-distances permettent un positionnement métrique lorsque l'on dispose de données différentielles. Or, la précision nécessaire sur la détermination des sommets de prises de vue doit être la même que celle que l'on veut au sol, c'est-à-dire $10 \mu\text{m}$ x l'échelle de prise de vue. Une précision décimétrique est donc nécessaire à la plupart des applications photogrammétriques (échelle variant de 1 : 5 000 à 1 : 40 000). Cette précision ne peut pas être atteinte avec l'utilisation exclusive des pseudo-distances. Mais la mesure de phase est ambiguë. En effet, en utilisant la phase, on ne mesure que la partie fractionnaire du cycle, le nombre entier de cycles séparant le récepteur du satellite est inconnu.

Il existe donc *a priori* deux manières d'améliorer la précision de la détermination des sommets de prises de vue : améliorer le traitement sur les pseudo-distances en utilisant la phase ou être capable de résoudre les ambiguïtés dans le traitement de la phase.

2.3.1.3.1. La trajectographie

Il s'agit d'un calcul en mode différentiel entre une station fixe de coordonnées parfaitement connues et le récepteur à bord de l'avion. Le procédé est basé sur le fait que les données sur la phase sont moins bruitées que les données de pseudo-distance. En effet, on considère que le bruit de la mesure est

estimé à mieux que 0.1 % de la longueur d'onde du signal sur lequel se fait la mesure. Par conséquent, le bruit sur les pseudo-distances est métrique tandis que le bruit sur la phase est millimétrique. On cherche par conséquent à améliorer le traitement des pseudo-distances en diminuant le bruit de mesure par un lissage à l'aide de la mesure sur la phase. Il s'agit du mode trajectographique. Dans ce mode, aucune initialisation n'est nécessaire. On obtient par cette méthode une cohérence interne sur un axe inférieur au décimètre. En revanche, la précision inter-axe n'est pas meilleure que quelques mètres.

2.3.1.3.2. La résolution d'ambiguïtés en vol

Là encore, il s'agit d'un positionnement en mode différentiel. Si dans le cas de deux récepteurs fixes, la résolution des ambiguïtés est possible sans trop de difficultés pour des lignes de base de plusieurs centaines de kilomètres, elle est, par contre, beaucoup plus difficile lorsqu'un des récepteurs est mobile.

Un algorithme développé par Remondi (1990), permet la résolution des ambiguïtés en vol (*On The Fly* ou OTF) sans initialisation en phase statique qui ne serait évidemment pas possible en avion.

L'initialisation de la fixation des ambiguïtés en vol ne peut se faire au-delà d'une distance de 10 km entre le récepteur fixe et le mobile. Une fois l'initialisation faite, la base peut aller jusqu'à 30 km, mais sans saut de cycle. Car s'il y a un saut de cycle, un logiciel OTF traditionnel reprendra l'initialisation, et donc le récepteur mobile devra se rapprocher à nouveau de la référence. Cette méthode de résolution des ambiguïtés en vol est répandue chez nombre de constructeurs et utilisateurs GPS, mais elle a l'inconvénient majeur, dans une application de production photogrammétrique, d'imposer des bases maximales d'une trentaine de kilomètres. Certains constructeurs vendent une solution pour des bases allant jusqu'à 50 km, mais les tests grandeur nature sont peu nombreux. Cela dit, l'ordre de grandeur reste le même et nécessite la présence d'un récepteur fixe proche de la zone à photographier, ce qui crée une contrainte opérationnelle très lourde, sauf si une station GPS permanente est présente.

Les positions GPS calculées au cours de la mission sont celles du centre de phase de l'antenne. Or, la position qui nous intéresse est celle du sommet de prise de vue au moment de la prise de vue. Cela nécessite une bonne synchronisation entre la prise de vue et le récepteur GPS et de connaître parfaitement le vecteur antenne caméra au moment de la prise de vue. Ce vecteur est constant dans le système avion, par contre il ne l'est pas dans le système terrain. Il faut donc calculer la matrice de passage entre les deux systèmes, soit en utilisant des points terrain, soit connaître au moment de la prise de vue l'attitude de l'avion.

2.3.1.4. *La mesure d'attitude*

La mesure d'attitude par GPS est une mesure interférométrique. On mesure les temps d'arrivée de l'onde sur un système à plusieurs antennes formant un solide indéformable. Cela nécessite une parfaite synchronisation sur les temps d'arrivée aux différentes antennes. Ceci est réalisé à l'aide de récepteurs, d'une horloge et de plusieurs antennes. On trouve actuellement sur le marché des systèmes à 3 ou 4 antennes.

Plus les antennes sont séparées l'une de l'autre, meilleure est la précision angulaire. Malheureusement, il est particulièrement difficile d'obtenir des polygones indéformables sur des avions servant à la prise de vue aérienne compte tenu de la flexibilité des ailes (à l'exception notable des avions à aile haute). Les précisions actuellement obtenues peuvent atteindre le millième de radian alors qu'une précision 10 à 100 fois supérieure serait nécessaire dans la plupart des applications de photogrammétrie.

Il est possible, mais à un coût élevé, de coupler au système GPS un système inertiel. L'utilisation du système inertiel seul n'est pas envisageable car il dérive très vite. En revanche, en recalant régulièrement la centrale à inertie avec les mesures GPS, il est possible d'obtenir une précision sur les mesures angulaires d'environ 10^{-5} radians. C'est cette configuration qui est employée avec les télémètres laser aéroportés.

Les informations d'attitude sont connues en temps réel dans l'avion et peuvent par conséquent être aussi utilisées pour la navigation afin de minimiser les erreurs sur le recouvrement et d'avoir un tableau d'assemblage suffisamment précis à la fin de la mission aérienne.

2.3.1.5. *Terminologie GPS*

Cinématique : méthode différentielle (DGPS) basée sur la mesure de phase d'au moins quatre satellites. Le principe consiste à résoudre les ambiguïtés entières par une initialisation, puis à stationner des points quelques secondes en conservant le signal sur les satellites durant les trajets, et donc les mêmes ambiguïtés entières.

Pseudo-cinématique : méthode cinématique originale qui consiste à restationner des points et à traiter le fichier d'observation comme si c'était une station fixe ayant une lacune dans ses acquisitions (les ambiguïtés étant les mêmes).

Dynamique : méthode différentielle (DGPS) avec observations de pseudo-distances, la station de référence (fixe) dont les coordonnées sont connues envoie par radio en temps réel des corrections à la station mobile, qui calcule alors sa position.

Trajectographie : mode spécifique qui utilise des observations de phase et pseudo-distances, le calcul se faisant avec les pseudo-distances lissées par la mesure de la phase de la porteuse.

Navigation : nous associerons à ce terme le sens d'aide à la navigation, c'est-à-dire la connaissance d'informations sur la position et l'attitude de l'avion qui pourraient permettre au pilote ainsi qu'aux opérateurs de rectifier leur trajectoire ou l'orientation de la caméra dans les situations le nécessitant.

2.3.2. Utilisation du GPS en aérotriangulation⁴

2.3.2.1. Comment limiter les travaux de stéréopréparation ?

Quel que soit le produit photogrammétrique à réaliser (restitution vectorielle, orthophoto, modèle numérique de terrain), il est indispensable de déterminer dans une phase préalable les formules d'image employées. Cette phase est traditionnellement appelée aérotriangulation (ou spatiotriangulation dans le cas d'images spatiales, de géométrie souvent différente). Nous rappellerons ici rapidement le principe de cette méthode classique en photogrammétrie analytique (on trouvera des études complètes de ce problème dans tout bon ouvrage de photogrammétrie analytique), et seuls seront développées les améliorations rendues possibles par l'apparition de la photogrammétrie numérique.

Rappelons que les seules propriétés géométriques des images ne permettent de connaître ni la position, ni l'orientation, ni l'échelle des objets photographiés. La détermination de la similitude mathématique correspondante ne peut être obtenue que par des mesures externes : ainsi, si on se limite au cas du couple photogrammétrique et à des déterminations de points connus sur le terrain, il faudra mesurer au minimum deux points planimétriques et trois points altimétriques.

L'aérotriangulation a pour but de réduire au maximum ce besoin en mesures terrain, en résolvant simultanément les formules d'image d'un grand nombre d'images. Il s'agit en fait d'ajuster au mieux (souvent au sens des moindres carrés) les mesures de coordonnées images de points de liaison homologues, un certain nombre de points terrain étant considérés comme connus, ainsi qu'éventuellement des mesures annexes enregistrées pendant le vol photographique, ou des données de trajectographie satellitale. On aboutit alors généralement à un système surabondant de plusieurs milliers d'équations non

4. Cette section a été rédigée par Yves EGELS et Michel KASSER.

linéaires, que l'on pourra résoudre par approximations successives à partir d'une solution approchée.

De nombreux logiciels du commerce permettent de mener à bien ce calcul. Sans mesures en vol, ils peuvent ramener les points terrain nécessaires à un point planimétrique pour 5 à 6 couples, et un point altimétrique par couple. La photogrammétrie numérique n'a pas apporté de différence de fond au niveau du calcul d'aérotriangulation lui-même, mais elle permet d'automatiser presque complètement la mesure des points de liaison, qui restait manuelle (et très fastidieuse) en analytique. De plus, le développement des caméras numériques, dans la mesure où elles nécessitent plus d'images pour couvrir une zone donnée, conduit à rechercher l'emploi de mesures annexes plus performantes afin de restreindre encore les mesures terrain.

Si le recouvrement entre bandes est très fort (par exemple 60 %), chaque point se retrouve dans trois bandes, mais au total, cela ne permet pas de diminuer sensiblement le nombre de points d'appui au-dessous de la moitié des valeurs pour une aérotriangulation normale. Par contre, on obtient ainsi une fiabilité bien plus grande, et la précision est elle aussi bien améliorée.

C'est une solution qui est employée pour le relevé de points qui doivent être mesurés avec beaucoup de précision, matérialisés alors par des cibles afin que le pointé soit idéal (mesure de déformations du sol). Mais le surcoût de cette solution n'en fait pas une solution rentable en matière de travaux cartographiques.

2.3.2.2. Utilisation de mesures annexes lors du vol

2.3.2.2.1. APR

On a ainsi utilisé des mesures d'APR (*Aerial Profile Recorder*, cf. section 2.8), méthode qui fournissait des distances verticales (radar ou laser) de l'avion vers le sol, avec comme référence altimétrique une surface isobare arbitraire prise comme référence. L'avion suivait des trajectoires formant une grille, les points d'impact du télémètre au sol étant contrôlés par les images acquises par une caméra. Cette méthodologie est tombée en désuétude, mais a rendu de grands services pour couvrir de vastes zones inhabitées (désert australien, Grand Nord au Canada).

2.3.2.2.2. Mesures de paramètres annexes lors de vols photographiques

Il faudrait idéalement parvenir à connaître pour chaque instant où une image est acquise d'une part la position du sommet de la gerbe perspective, c'est-à-dire ici du centre optique de la chambre de prise de vues, dans le référentiel du chantier et d'autre part l'orientation de la chambre.

Ces deux groupes de paramètres ne sont pas du tout équivalents en termes de coût des mesures nécessaires à leur acquisition :

- pour la position du sommet, notons que l'on n'emploie pratiquement jamais le baromètre différentiel, qui pourtant pourrait apporter des éléments très intéressants, en créant des contraintes nouvelles fortes sur ce qui est une des faiblesses majeures de l'aérotriangulation, c'est-à-dire l'altimétrie. Cet équipement est par ailleurs fonctionnel sans aucune infrastructure externe, au contraire du GPS que nous allons examiner, autre solution plus complète au demeurant ;

- pour l'orientation de la chambre, il n'existe pas de capteur fiable et bon marché. Les centrales inertielles disponibles sont onéreuses. Basées sur l'emploi d'accéléromètres dont les valeurs sont intégrées deux fois pour obtenir les déplacements, et de gyromètres ou de gyroscopes pour mesurer les variations d'attitude, elles ont par nature même une dérive sensiblement proportionnelle au carré du temps, ce qui exige de les recalculer assez fréquemment, ce recalage ne peut alors être apporté que par le GPS. Il existe également des solutions utilisant le GPS seul, en mesurant simultanément les données GPS acquises sur trois ou quatre antennes de façon synchrone. Il s'agit alors d'un mode qui s'apparente à de l'interférométrie. Sa sensibilité angulaire dépend en fait de l'écartement des antennes, les valeurs de mesure de la composante verticale n'étant guère meilleures que le centimètre. Ceci implique, les antennes ne pouvant pas être plus distantes que 10 m (antennes sur les ailes, le nez et l'aileron ou la dérive de l'avion) une précision de mesure qui n'est pas supérieure à quelques mrd, avec tous les aléas de la mesure GPS. Par comparaison, avec un pointé sur cliché à 15 μm près (ce qui n'a rien d'excellent) et une focale classique de 150 mm, on comprend que le besoin en orientation absolue est compris entre 10^{-4} et 10^{-5} rd, ce qui disqualifie de fait le GPS employé ainsi. Par ailleurs, il faudrait que les quatre antennes forment un polyèdre indéformable ; ceci n'est pas le cas, puisque la plupart des ailes d'avions fléchissent en fonction des contraintes apportées par les turbulences : la seule exception notable est formée par les avions à aile haute, lorsque l'implantation des antennes sur les ailes est localisée sur la zone d'accrochage des haubans qui retiennent l'aile.

2.3.2.3. Positionnement par GPS

La précision qu'il faut atteindre pour localiser les sommets doit être au moins aussi bonne que celle des points à localiser au sol, et dépend évidemment des paramètres de la prise de vue (échelle, hauteur de vol, taille de pixel, etc.). Dans les cas courants, il s'agit d'une précision allant de 5 à 30 cm. Nous sommes dans un régime dynamique, avec une vitesse de l'avion de 100 à 200 m/s. Ceci implique une très bonne synchronisation entre la prise de vue et le récepteur GPS, ce qui ne pose pas de problèmes particuliers avec les

matériels GPS disposant d'une entrée adaptée, la datation d'évènements externes étant couramment pratiquée avec une précision qui atteint la μ s. Mais, bien entendu, il faut que la chambre de prises de vues mette à disposition un signal parfaitement synchrone (à 0.1 ms près) de l'ouverture de l'obturateur, ce qui est le cas avec toutes les chambres modernes mais n'est pas toujours vrai sur des matériels anciens.

En plus il faut obtenir un positionnement continu au cours du mouvement. Avec des mesures GPS toutes les secondes, on est amené à interpoler la position de la chambre lors du signal de synchronisation sur des distances de 100 à 200 mètres, ce qui évidemment ne permet pas une précision décimétrique dès que l'atmosphère est un peu turbulente. Et il faut bien noter que plus le besoin de précision est grand, plus on vole à altitude basse et donc plus la turbulence est forte. Donc un positionnement adapté à des grandes échelles exige un échantillonnage temporel meilleur que 1 s, une valeur de 0.1 s (pourtant rarement disponible sur les récepteurs GPS) étant même à peine suffisante. Dans cette optique, on peut associer un récepteur GPS et une centrale inertielle qui sert alors d'interpolateur précis, la dérive de la centrale étant faible sur une seconde, et la centrale pouvant fournir des données très rapprochées dans le temps (typ. de 100 à 1000 Hz). Mais cette solution est assez coûteuse et encore peu employée.

Comme nous l'avons vu au paragraphe 2.3.1, les besoins de précision requis exigent l'emploi du GPS en mode différentiel. Deux modes peuvent être pratiqués :

- mode cinématique, avec mesure de la phase sur la porteuse du signal. Soit on initialise avec une durée permettant la résolution des ambiguïtés (ceci exige quelques minutes), et on les fixe alors pour le reste du vol, mais c'est peu fiable et ce n'est plus guère employé car, lorsque l'avion tourne, il arrive fréquemment qu'il y ait de courtes interruptions du signal, ruptures dans la continuité de la mesure qui obligerait alors à résoudre de nouveau les ambiguïtés, ce qui évidemment devient impossible en vol. Soit on résout les ambiguïtés au vol (méthode OTF, *on the fly*), ce qui corrige cette difficulté. Néanmoins, il faut bien noter que jusqu'ici cette méthode ne fonctionne qu'avec des bases assez courtes (20 ou 30 km au maximum), ce qui est peu compatible avec les conditions de travail des avions photographes. En effet, il faudrait, avant le commencement du vol, mettre en place une station de réception GPS à proximité de la zone à photographier. Les aléas de programmation des vols photographiques dans les pays tempérés rendent une telle démarche assez contraignante, la météo étant à elle seule une contrainte suffisamment forte pour ne pas rajouter une obligation de mise en place d'un tel matériel, ce qui ferait perdre trop de temps. Les seuls cas où l'on puisse travailler selon ce mode

seront ceux où existe une station GPS permanente à proximité de la zone à photographier. Cette méthode est donc assez peu employée ;

– mode trajectographie, on mesure alors les pseudo-distances ainsi que la phase. La précision est moindre, mais on peut travailler avec des distances à la station de référence beaucoup plus importantes. Nous allons étudier plus en détail cette solution.

2.3.2.4. Étude du mode trajectographie

La distance satellite-récepteur peut être d'une part calculée à partir du temps de parcours de l'onde déduite de la datation du code C/A, de la correction d'horloge, de la correction troposphérique et de la correction ionosphérique (précision typique de l'ordre d'un mètre en mode différentiel). Elle peut être d'autre part calculée à partir de la mesure de phase, d'un nombre entier k de tours de phase, de la correction d'horloge, de la correction troposphérique et de la correction ionosphérique (la précision est alors millimétrique, sauf que k est inconnu). La différence entre ces deux variantes de la même distance étant nécessairement nulle en théorie, en intégrant les mesures sur un satellite donné pendant un temps suffisant on peut obtenir une estimation de k . Ce mode est assez peu adapté à des opérations temps réel, mais tant qu'il y a des satellites en visibilité commune entre la station fixe et l'avion, le résultat est utilisable pour autant que le GDOP soit correct, les bases traitées pouvant dépasser le millier de kilomètres.

Les résultats sont en général toujours du même type : sur quelques minutes le bruit de détermination des coordonnées de l'avion est inférieur à 5 cm, et on observe des biais, erreurs de positionnement variant lentement, avec une constante de temps typique de 1 à 2 heures, et des amplitudes de l'erreur de l'ordre du mètre. Ceci est dû aux changements lents de la configuration des satellites dans le ciel, les erreurs d'orbites se répercutant directement dans le calcul du point sous des formes variant sans cesse.

On modélisera alors les mesures GPS sous forme d'une erreur de translation générale des positions GPS pour chaque tranche temporelle courte pertinente, typiquement de 10 mn, par exemple pour chaque série de photographies le long d'un axe de vol. On rajoute ainsi un triplet d'inconnues supplémentaires dans l'aérotriangulation pour chaque bande. Si l'on raisonne en différentiel, la précision est alors centimétrique, et en mode absolu c'est sans importance puisqu'il y a des points d'appui au sol qui imposent alors complètement la translation voulue pour l'ensemble des clichés.

On parvient ainsi à diminuer encore considérablement le besoin de points au sol : En planimétrie, il en faut au minimum 4 par bloc, plus évidemment

quelques-uns pour contrôles, et en altimétrie 1 point tous les 20 ou 30 couples est satisfaisant. Ces points n'ont pas d'obligation absolue d'être aux bordures extrêmes du chantier.

Il convient toutefois d'être très attentif aux problèmes qui sont alors posés par les systèmes de référence employés. En particulier en altimétrie, rappelons que les mesures GPS sont purement géométriques, alors que le nivellement de référence est toujours basé sur des mesures géopotentielles. Les pentes du géoïde, ou plutôt de la surface d'altitude nulle, doivent alors être prises en compte selon la précision demandée au lever.

On peut envisager de traiter le chantier sans aucun point de référence au sol (zones inaccessibles par exemple). Dans ce cas, on obtient un lever cohérent mais qui est à une translation près d'ordre de grandeur métrique, translation qui est en gros la moyenne des inconnues de systématisme.

Il convient également d'être attentif à la géométrie de l'ensemble des points de stéréopréparation. S'ils sont tous à peu près alignés, le basculement transverse sera indéterminé. Il faut donc inclure des points dans les zones de recouvrement latérales entre bandes. Si l'on travaille sur une zone sans équipement en nivellement suffisant, on pourra améliorer la qualité de l'aérotriangulation en recoupant les bandes parallèles par des bandes de photos transverses. Mais il pourra être nécessaire de faire du nivellement sur le terrain si le géoïde est inconnu, uniquement pour en imposer la pente dans le modèle du chantier.

2.3.2.5. *Raccordement de l'antenne GPS au centre optique de la chambre de prise de vues*

On commencera par mesurer avec des méthodes classiques de topométrie le vecteur d'excentrement E allant de l'antenne au centre optique, dans le référentiel de l'avion. M_A étant la position de l'antenne, M_S celle du centre optique, on pourra écrire :

$$X_S = X_A + T + R \cdot E \quad [2.14]$$

où V est le vecteur décrivant le systématisme résiduel de ces mesures, et R décrit la rotation de l'avion dans le référentiel du chantier, supposée égale à la rotation de la chambre, elle même déduite de la rotation du faisceau, sous-produit de l'équation de colinéarité.

2.3.2.6. *Conclusion*

L'utilisation du GPS (ou de n'importe quel autre GNSS système positionnement global, comme le GLONASS russe ou le futur GALILEO

européen) dans l'avion est une source de données auxiliaires très importante pour réduire les coûts engendrés par la mesure de points d'appui nécessaires à l'aérotriangulation. Mais dans de tels cas, il faut aussi bien prendre en compte les imperfections du GPS, et savoir expertiser les mesures à prendre quand des ensembles de données sont inexploitable. D'un autre côté, il faut bien comprendre que si des mesures inertielles sont susceptibles de fournir une excellente interpolation dans les mesures GPS, elles ne peuvent corriger leurs éventuels défauts, et de telles données doivent être comprises comme auxiliaires. Tout l'effort possible doit être fait pour obtenir le meilleur signal dans l'antenne : un dispositif permettant l'orientation absolue de la chambre de prises de vues sera toujours bienvenu, mais en attendant sa venue éventuelle, le GPS donne dès à présent d'excellents résultats.

2.4. Automatisation de l'aérotriangulation⁵

2.4.1. Introduction

2.4.1.1. Présentation

Cette partie concerne la détermination automatique de points de liaison et leur emploi dans deux domaines : aérotriangulation et réalisation automatique de tableaux d'assemblage. Le lecteur portera son attention sur le fait qu'on ne considère pas ici la mesure des points d'appui. On ne traitera ni de leur détermination (qui est faite par des techniques de géodésie), ni de la mesure de leur position dans les images (qui sont mesurées manuellement).

Le calcul automatique de points de liaison est un sujet d'intérêt grandissant notamment en raison du nombre croissant d'images produites pour des chantiers photogrammétriques. En particulier, l'usage qui se répand maintenant de caméras numériques d'une part, et de forts recouvrements (notamment interbandes) d'autre part, engendre des images en nombre plus important que par le passé. Il est désormais possible de trouver un nombre appréciable de fabricants de logiciels d'aérotriangulation proposant un module de détermination automatique de points de liaison. Ces modules sont en mesure de donner de bons résultats sur des scènes simples (ne comportant pas trop de relief ni de zones texturées trop importantes). Par ailleurs, la détermination des points de liaison dans des configurations moins simples mobilise encore grandement la communauté des chercheurs. Ces configurations moins simples sont notamment les chantiers photogrammétriques terrestres, les prises de vues aériennes à dates différentes, les scènes à fort relief, etc. Ce type de problématique fait l'objet d'un suivi de la part de l'OEEPE [HEI 99].

5. Cette section a été rédigée par Franck JUNG, Franck FUCHS et Didier BOLDO.

2.4.1.2. *Notions de base*

Il convient de rappeler quelques notions qui seront utilisées dans la suite de cette partie :

- point de liaison : triplet de coordonnées 3D correspondant souvent à la position d'un détail physique de la scène, et vu dans au moins deux images ;

- mesure : coordonnées 2D de la projection d'un point (d'appui, ou de liaison) dans une image ;

- un point d'intérêt est un point de l'image autour duquel le signal présente des propriétés particulières, notamment celle de posséder des dérivées fortes dans plusieurs directions ou au moins dans deux directions orthogonales, détecté par un outil théorique particulier (exemple/ détecteur de coins, de jonction...).
Mesure de ressemblance : fonction associant à deux voisinages de deux points une valeur réelle, généralement bornée. La mesure la plus souvent utilisée est le coefficient de corrélation linéaire ;

- points homologues : ensemble de points satisfaisant certaines propriétés au regard de la fonction de ressemblance. Généralement, un point P_1 d'une image I_1 et un point P_2 d'une image I_2 sont jugés homologues si pour tout P de I_1 , P_1 et P_2 sont plus ressemblants que P et P_2 , et pour tout Q de I_2 , P_1 et P_2 sont plus ressemblants que P_1 et Q ;

- répétabilité : qualité d'un détecteur de points d'intérêt capable, pour une scène donnée, de détecter les points d'intérêt correspondant aux mêmes détails dans des images différentes, même si les conditions de prise de vue varient (éclairage, point de vue, échelle, rotation...). Un détecteur de points d'intérêt est d'autant plus répétable qu'il produit les mêmes ensembles de points pour une même scène malgré les variations de condition de prise de vue. Cette notion est précisée dans [SCH 96] ;

- disparité : pour deux images possédant un angle de rotation faible l'une par rapport à l'autre, la disparité de deux points est le vecteur égal à la différence de leurs coordonnées. Dans le cas de prises de vues aériennes avec visée verticale, la disparité de deux points correspondant au même détail physique du terrain est directement liée à l'altitude de ce détail. Cette notion se généralise au cas d'images possédant une rotation relative quelconque.

2.4.2. *Objectifs de la sélection de points d'intérêt*

Cette section discute trois points concernant les objectifs des méthodes de mesure automatique de points de liaison. On traitera en premier lieu de leur fiabilité, puis de leur précision, ainsi que d'un troisième point important, qui place la mesure automatique de points de liaison en net décalage par rapport aux techniques manuelles : il s'agit du nombre de points de liaison.

2.4.2.1. *Fiabilité*

La plupart des méthodes d'aérotriangulation s'appuient sur une optimisation en utilisant la technique des moindres carrés. Cette technique est particulièrement sensible aux valeurs aberrantes. Il est donc nécessaire, lors du calcul des points de liaison, de ne pas fournir de mesures aberrantes.

Le premier objectif majeur que doit donc atteindre une méthode de mesure automatique de points de liaison est de fournir des points qui ne soient pas entachés de fautes.

2.4.2.2. *Précision*

Ce second objectif majeur peut être l'objet de travaux approfondis à partir du moment où l'on sait que le chantier d'aérotriangulation a été mis en place de façon correcte (*i.e.* il n'y a plus de fautes).

L'obtention d'une meilleure précision pour l'aérotriangulation passe par la mesure plus précise des points de liaison, mais aussi par la prise en compte, dans les méthodes d'aérotriangulation, d'un modèle d'erreur adapté aux véritables erreurs faites par les outils de mesure (ce qui implique notamment d'étudier effectivement ces erreurs pour les modéliser). En effet, les modèles d'erreur utilisés sont généralement gaussiens, et il reste à établir que ce modèle est pertinent pour tel ou tel outil de mesure. Le cas échéant, une adaptation du modèle d'erreur resterait à faire.

2.4.2.3. *Nombre de points de liaison*

L'aérotriangulation traditionnelle utilise 15 mesures par image, alors que les outils de mesure automatique de points de liaison sont rapidement capables de fournir de très nombreuses données.

Dans le cas manuel, ces points sont très pertinents, car sélectionnés par un opérateur. Leur répartition géométrique peut notamment être assurée *a priori* par l'opérateur. Dans le cas automatique, le manque d'intelligence des méthodes informatiques est compensé par l'abondance de données. Concernant la répartition des points d'une part, il est difficile de forcer la machine à « trouver » une solution dans une région donnée, et il est donc probablement préférable de lui laisser trouver une autre solution au voisinage de la zone souhaitée. Concernant les fautes, l'abondance de données ne réduit pas les taux de fautes des méthodes, mais permet la mise en œuvre pertinente d'outils statistiques visant à éliminer ces fautes.

Les études ne sont toutefois pas suffisamment avancées pour permettre de fixer une valeur idéale de mesure par cliché. On s'attend tout de même à ce que

le nombre de points requis soit sensiblement plus élevé dans le cas automatique (typiquement 100 mesures par cliché, [HEI 98]). Le débat reste très ouvert sur cette question.

2.4.3. Méthodes

La mesure automatique de points de liaisons fait généralement appel à deux étapes majeures. La première concerne la détection et la localisation de points d'intérêt dans les images. Les points d'intérêt peuvent être calculés à partir des images prises individuellement. Ensuite, ces points sont utilisés pour produire les points de liaison proprement dits. Le passage des points d'intérêt aux points de liaison peut se faire généralement suivant deux stratégies opposées.

Dans les deux cas, les stratégies recherchent entre les images des points pour lesquels la valeur de mesure de ressemblance utilisée est maximale. Les stratégies diffèrent par l'espace de recherche du maximum.

Pour un point d'intérêt P d'une image I, la première stratégie consiste à rechercher dans une image J, la position du point Q qui maximise la valeur de ressemblance parmi toutes les positions possibles (tous les pixels d'une région où l'on s'attend à trouver Q). On peut même envisager de calculer la position de Q avec une précision sub-pixellaire. Cette stratégie, que l'on appellera « radiométrique », est orientée : I ne joue pas le même rôle que J dans ce cas.

La seconde stratégie consiste à restreindre l'espace de recherche dans J aux seuls points d'intérêt déjà calculés. On appellera cette approche « géométrique » car elle se focalise sur la position des points de liaison.

2.4.3.1. Méthodes de détection de points

Les détecteurs de points utilisent des caractéristiques particulières du signal. Nous citerons plus particulièrement le détecteur de Förstner [FÖR 87] (très utilisé dans les logiciels du commerce) ainsi que le détecteur de Harris qui montre une bonne répétabilité [HAR 88], [SCH 96].

Ces détecteurs ont généralement une mauvaise précision de localisation. Des expériences ont montré un décalage (emq) de 1 à 1.5 pixel entre les coins réels et les points détectés [FLA 99]. Néanmoins, la délocalisation de ces points possède un très fort systématisme et un faible écart type (voisin de 1/2 pixel). De ce fait, il est possible de considérer ces points satisfaisants pour une mise en place photogrammétrique. Une meilleure localisation de ces points détectés peut être envisagée par des techniques spécifiques (cf. paragraphe 2.4.3.4).

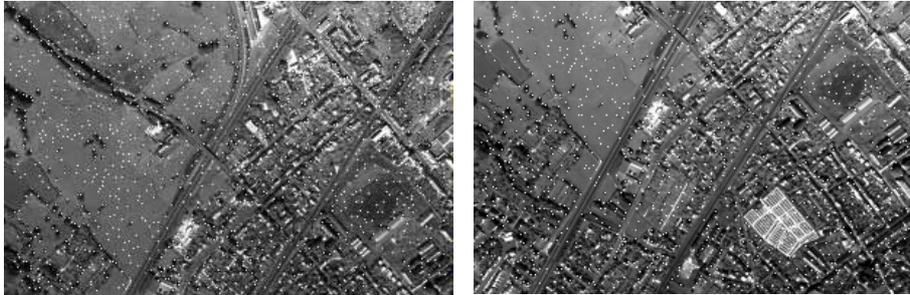


Figure 2.27. Deux images 300 x 200 et leurs points d'intérêt

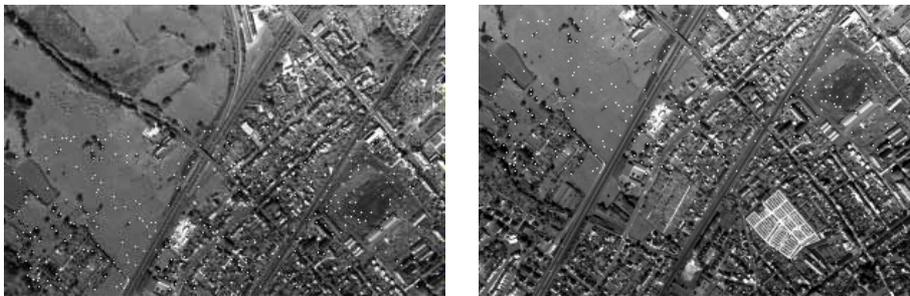


Figure 2.28. Résultat de mise en correspondance

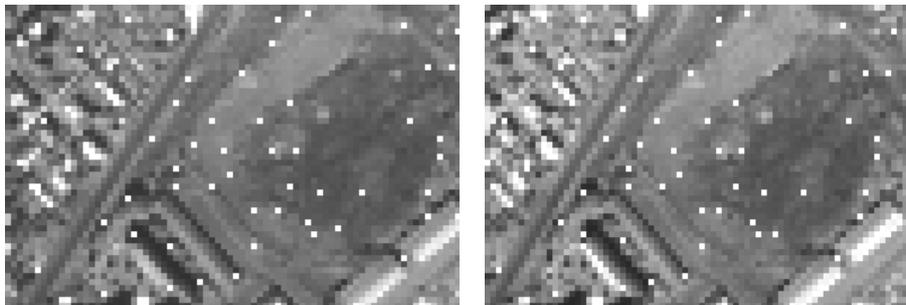


Figure 2.29. Extrait des images avec les points homologues

La figure 2.27 présente deux images voisines (taille : 300 x 200) avec, en blanc, leurs points d'intérêt. Étant donné la densité des points, leur représentation est difficile, et l'étude visuelle des images doit être minutieuse. La figure 2.28 présente les points mis en correspondance dans ces deux images.

Une étude attentive permet de voir que les appariements sont très souvent corrects. Il est significatif de constater que les points homologues se situent dans la zone de recouvrement des deux images. La figure 2.29 présente deux extraits de ces images, permettant de voir en détail les points mis en correspondance dans une zone localisée des images.

2.4.3.2. *Calculs de points homologues*

Aspect photométrique : un critère de ressemblance très utilisé dans de nombreuses techniques d'appariement se trouve être le coefficient de corrélation. Ce coefficient est calculé sur deux vignettes. L'avantage de cette technique réside dans le fait de pouvoir appairer deux zones ayant des niveaux de gris ressemblants. Néanmoins, il est nécessaire d'utiliser des images parfaitement orientées (problème des images comportant de fortes rotations).

Aspect signal : un critère uniquement basé sur la ressemblance locale (critère photométrique) n'est pas toujours suffisant. En effet, la corrélation au sein d'une zone homogène ou le long d'un contour reste une source d'ambiguïté. Une stratégie concurrente ou complémentaire consiste donc à n'appairer que des points possédant des caractéristiques particulières du signal.

Une approche mixte peut permettre de tirer pleinement avantage des deux méthodes. La détection des points se fait suivant des caractéristiques du signal et l'appariement est effectué suivant un critère photométrique.

2.4.3.3. *Fiabilité*

Les outils contribuant à l'obtention d'une bonne fiabilité des points de liaison sont assez nombreux :

- répétabilité du détecteur : on a intérêt à utiliser un détecteur de point possédant une bonne répétabilité. On citera les détecteurs de Förstner [FÖR 87] et Harris [HAR 88]), dont Schmid [SCH 96] montre que celui de Harris possède les meilleures propriétés de répétitivité ;

- approche multi-échelle : cette approche permet de guider la recherche du point homologue en pleine résolution à l'aide d'une prédiction fiable de la disparité sur des images sous-échantillonnées. Cette approche permet de résoudre deux problèmes : d'une part des problèmes de fautes très grossières (on peut volontairement limiter la taille des zones de recherche) ; d'autre part cette approche permet de limiter sensiblement la combinatoire du problème de mise en correspondance à pleine échelle. On peut noter que l'introduction d'un modèle numérique de terrain peut remplacer partiellement l'utilisation d'images sous-échantillonnées pour prédire les zones de recherche en pleine échelle ;

– multiplicité : une évaluation visuelle de la qualité des points de liaison permet de constater que le pourcentage de points de liaison faux diminue sensiblement avec l'ordre du point (nombre d'images pour lesquelles le point est visible). Il y a néanmoins plusieurs façons d'exploiter cette constatation. Une façon très fiable est de considérer un point multiple valide si toutes les mesures associées sont en interconnexion totale au regard de la fonction de ressemblance, c'est-à-dire si chaque paire de mesure est reconnue valide par un processus de mise en correspondance image par image. La figure 2.30 illustre ce mécanisme : chaque couple de points de cette liaison multiple a été détecté comme paire de points homologues par l'algorithme d'appariement.

Les points de liaison multiples possédant une interconnexion partielle, le premier ensemble est bien évidemment inclus dans le second. Selon le nombre de points multiples obtenus, il sera possible de ne prendre que des points de liaison du premier ensemble (qui sont *a priori* plus fiables) ou il sera nécessaire d'ajouter un certain nombre de points appartenant au second ensemble et non au premier. Il faut également noter que les points à forte multiplicité sont nécessaires pour le calcul de l'aérotriangulation.

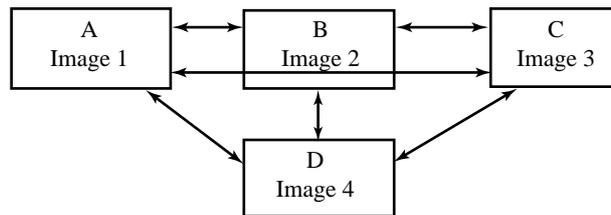


Figure 2.30. Exemple de point multiple d'ordre 4 à interconnexion totale

La figure 2.32 présente un exemple de résultat de mesure de points multiple. Son ordre est 4. Il est issu d'une prise de vues sur Le Mans, avec une caméra numérique de l'IGN.

Filtrage photogrammétrique : en dehors de toute considération de type radiométrique, on peut également utiliser les propriétés photogrammétriques des points de liaison. En effet, ces points sont censés représenter un même point terrain. Par conséquent, tous les rayons perspectifs doivent se couper en un même point. Plusieurs méthodes peuvent être proposées pour tirer parti de cette propriété. La première consiste à tenter de mettre en place chaque couple individuellement, par exemple par un algorithme de type 8 points [HAR 97]. Techniquement, cet algorithme calcule la matrice fondamentale associée à un couple de clichés. Si on note \mathbf{u} et \mathbf{u}' les vecteurs projectifs représentant, dans le repère cliché, deux points homologues, la matrice fondamentale F est définie

par : $\mathbf{u}'^T F \mathbf{u} = 0$. Moyennant quelques précautions numériques, l'algorithme permet de calculer la matrice fondamentale la plus probable associée à un échantillon de points. À partir de la matrice fondamentale, il est possible de déterminer la droite épipolaire associée à un point \mathbf{u} , et donc également de calculer la distance entre \mathbf{u}' , homologue de \mathbf{u} , et cette droite épipolaire. Cette distance est une forme de « résidu » et permet donc de qualifier les points. Cet algorithme est très général, et permet une mise en place quelle que soit la configuration. Il a donc été utilisé pour tous les ensembles de deux clichés ayant au moins neuf points en commun, quelles que soient leurs positions respectives. Dans la pratique, pour chaque « couple » on tente une mise en correspondance itérative : on met en place, on calcule les résidus, on élimine les points ayant un résidu supérieur à trois EMQ, et on réitère la mise en place, et ce jusqu'à ce que celle-ci soit satisfaisante, ou que le procédé soit stabilisé. Si la convergence n'a pu avoir lieu, on essaie alors un algorithme aléatoire [ZHA 96] : on choisit N échantillons de points de liaison et on applique l'algorithme sur chaque échantillon. Les résidus sont ensuite calculés sur l'ensemble des points, et on conserve la mise en place dont la médiane des résidus est la plus faible. Actuellement, on effectue 200 tirages de 16 points, et les résultats semblent à peu près constants. Si l'une des méthodes a donné une mise en place satisfaisante, on élimine tous les points dont le résidu est supérieur à trois fois la médiane.

Si le nombre de points justes excède sensiblement le nombre de fautes, la mise en place est valable, et les points dont l'écart avec la mise en place est trop important sont éliminés. Une autre méthode est d'utiliser réellement la photogrammétrie. En effectuant une mise en place approchée, les fautes apparaissent. La seule limitation est que la plupart des logiciels d'aérotriangulation ne sont pas prévus pour gérer les milliers de points générés automatiquement par les algorithmes. Des outils adaptés doivent donc être mis en place.

À l'issue du filtrage, on souhaite généralement choisir parmi les points restants, en fonction de critères tels que la répartition ou la validité *a priori*. En effet, la répartition des points n'est absolument pas uniforme, certaines zones étant suréquipées par rapport à d'autres. Afin de limiter ce phénomène et de limiter le nombre de points dans l'aérotriangulation, il est nécessaire de choisir parmi l'ensemble de points filtrés. Les critères à prendre en compte sont l'équité répartition des points, la validité *a priori* de ceux-ci (en général, on peut évaluer la validité des points pendant leur calcul), afin de garantir d'avoir des points justes dans chaque relation interclichés.



Figure 2.31 a. Exemples de sources de fautes (en haut : véhicule en mouvement, à droite : zones ressemblantes localement)



Figure 2.31 b. Exemple de source d'erreur : zone texturée



Figure 2.31 c. Exemple de point correct

2.4.3.4. Précision

Pour l'obtention de points de liaison précis, deux grandes pistes existent. La première consiste à tenter de rechercher des positions sub-pixellaires qui maximisent la fonction de ressemblance. Cette technique s'apparente fortement aux méthodes d'optimisation. Le point important concernant cette méthode est qu'elle lie fortement précision et ressemblance : l'outil visant à mettre en correspondance les points et à les localiser de façon précise est le même.

La seconde est plus géométrique et découple les questions de ressemblance et de précision. Elle vise à déterminer des points correspondant à certains détails du terrain de façon précise, sans chercher à les mettre *simultanément* en correspondance (on notera toutefois que l'étape de détection n'est pas totalement découplée de la mise en correspondance. Il faut en effet veiller à chercher des points susceptibles d'être bien mis en correspondance, notamment des points possédant de bonnes propriétés géométriques, par exemple les coins, les jonctions...). Une démarche possible consiste à utiliser des modèles théoriques de coins ou positionner les points sur des intersections de lignes de contours déterminées précisément. La mise en correspondance de ces points ne remettra pas en cause leur localisation.

De façon générale, la précision de localisation des mesures dans les images dépend de l'algorithme de détection de points de liaison utilisé. L'impact de points d'intérêt ayant une précision subpixellaire sur la qualité de l'aérotriangulation reste dans une très large mesure à établir. L'aspect appariement dans ce type de démarche reste également très ardu. Les coordonnées subpixellaires des points homologues peuvent être estimées séparément ou bien conjointement à la phase d'appariement.

Pour cette approche, on peut citer notamment les techniques à base de modèles théoriques de coins, ou Blaszkla [BLA 94], ou Flandin [FLAN 99].

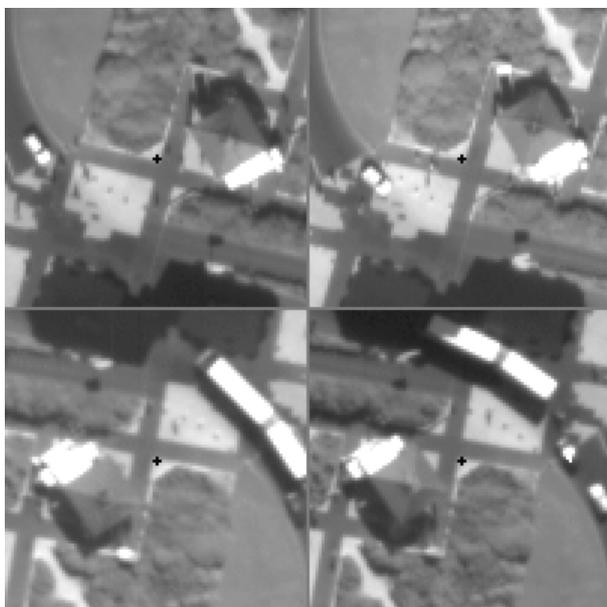


Figure 2.32. Exemple de point multiple d'ordre 4. Les images de la seconde ligne sont inversées car elles sont dans une bande différente

2.4.3.5. Place de la photogrammétrie

Même si le problème se place en amont de l'aérotriangulation, la photogrammétrie peut constituer une aide considérable pour résoudre le problème de la mesure des points de liaison.

Nous verrons dans la suite que la photogrammétrie peut guider les méthodes (par une approche multi-échelle), réduire la combinatoire du problème (par la prédiction des positions des homologues lors de l'appariement), et enfin filtrer les points de liaison après leur calcul, ce qui contribue sensiblement à leur fiabilité.

2.4.4. Discussion

Les premières parties ont exposé des considérations générales sur les objectifs et les méthodes liées à la mesure automatique de points de liaison. Cette partie discutera les limites et extensions des processus existants. En effet, même si les systèmes existants donnent de bons résultats, il existe de nombreux cas problématiques, ainsi que de nombreuses extensions possibles aux cas

moins classiques. Ces problèmes se décomposent suivant plusieurs axes intervenant dans les chantiers photogrammétriques : le capteur, la scène, la prise de vue.

Enfin, on présentera quelques considérations algorithmiques.

2.4.4.1. *Le capteur*

2.4.4.1.1. Utilisation de la couleur

A priori, l'utilisation de la couleur doit permettre une meilleure identification des détails du terrain, mais il s'avère que, actuellement, les méthodes mises en œuvre utilisent principalement des images en niveaux de gris, autant pour la détection de points d'intérêt que pour le calcul de points de liaison.

2.4.4.1.2. Qualité du signal

On s'attend *a priori* à obtenir des points de liaison de meilleure qualité avec un capteur possédant de meilleures propriétés : les détecteurs de points d'intérêt utilisent généralement les propriétés différentielles du signal, ils sont donc sensibles au bruit des images. Ainsi, un meilleur rapport signal sur bruit dans les images entraîne une plus grande stabilité de ces opérateurs. Dans la pratique, l'évaluation de la sensibilité au capteur en situation réelle nécessite de mettre en œuvre des matériels importants pour effectuer des vols dans des conditions identiques, si bien qu'aucune étude n'a été menée sur la question.

2.4.4.1.3. Données multicapteurs

L'intégration de données multicapteurs est un problème qui surviendra, car on peut envisager d'embarquer simultanément plusieurs caméras pour un vol (par exemple, embarquement simultané de capteurs spécialisés par canal : rouge, vert, bleu, plus infrarouge...). Aucune étude n'est disponible sur cette question.

2.4.4.2. *La scène*

2.4.4.2.1. Variation de l'aspect des objets

L'aspect des objets peut varier en fonction du point de vue. En cas de pente, les déformations perspectives peuvent être sensibles (le pire des cas est celui des façades en visée verticale). Les occlusions ne sont pas les mêmes. Enfin, l'aspect peut changer en fonction du point de vue. La figure 2.33 présente deux images d'une même scène, prises aux cours d'un même vol, de deux points de vue différents. La superposition des deux images montre nettement les variations d'intensité vues depuis ces deux points de vue.



Figure 2.33. *Exemple de changement d'aspect en fonction du point de vue*

2.4.4.2.2. Forêts, ombres, textures

Ces objets posent généralement des problèmes en analyse d'images. Dans le cas de la mesure de points de liaison, ils provoquent trois types de problèmes. Les arbres posent des problèmes de précision de pointé. Les ombres bougent entre deux prises de vues, notamment entre deux bandes (l'intervalle de temps entre deux images voisines de deux bandes différentes peut excéder 1 h), et donc, même si les extrémités des ombres peuvent sembler faciles à pointer, elles ne sont pas valides d'un point de vue photogrammétrique. Les textures posent naturellement des problèmes d'identification : il est facile de confondre deux détails très ressemblants dans des zones à motif répétitif, typiquement : les passages pour piétons.



Figure 2.34. *Exemple de point extrémité d'ombre*

2.4.4.2.3. Relief

Le relief provoque notamment des changements d'échelle et des variations de disparité. Lors de la recherche de points homologues, l'espace de recherche est généralement plus grand. Dans ce cas, l'utilisation de la photogrammétrie sur les images sous-échantillonnées (approche multi-échelle) peut s'avérer utile.

2.4.4.2.4. Cas de la photogrammétrie terrestre



Figure 2.35. *Exemple de prise de vue terrestre : les déformations perspectives peuvent être importantes*

Dans ce cas, bien souvent l'hypothèse d'un terrain plan (et horizontal) ne tient plus. L'espace des disparités est plus complexe que dans le cas aérien (dans le cas aérien, avec peu de relief, deux images d'une même scène sont en effet presque superposables, ce qui n'est pas le cas en imagerie terrestre). La figure 2.35 présente le cas d'une scène terrestre où une rue débouche suivant l'axe optique : les déformations liées à la perspective sont très importantes, ainsi que les disparités des points.

2.4.4.2.5. La prise de vue

Rotation des images

L'intérêt d'utiliser des images de rotation quelconque est de réaliser des bandes diagonales dans les blocs photogrammétriques, ce qui les rend plus rigides. Par ailleurs, si on cherche à calculer simultanément deux chantiers photogrammétriques se recouvrant, le cas de rotations quelconques peut se produire. Les techniques de mise en correspondance utilisent généralement des coefficients de corrélation linéaire entre fenêtres fixes. En cas de rotation connue entre les images, celles-ci peuvent être rééchantillonnées de telle sorte que ces techniques fonctionnent. En cas de rotation d'angle inconnu, il faut adapter les méthodes. Des mesures de ressemblances invariantes par rotation ont été développées : on peut envisager de tourner les vignettes [HSI 97], ou considérer des mesures s'appuyant directement sur des invariants [SCH 96]).

Prises de vues à diverses échelles

Dans le cas de prises de vues à différentes échelles, les mesures de ressemblances doivent être adaptées également, car elles ne supportent généralement pas les forts changements d'échelle. On connaît toutefois les échelles de prise de vue, ce qui réduit le problème. Mais même à échelles relatives connues, le problème est important, car pour comparer deux imagerie à deux échelles différentes, il faut adapter l'un des signaux. Par ailleurs, pour le positionnement des points d'intérêt par exemple, les erreurs commises par les détecteurs peuvent être différentes entre deux échelles.

Prises de vues à des dates différentes

Si l'on souhaite traiter simultanément des images prises à des dates différentes, il faut prendre garde notamment aux inversions de contraste qui se produisent (surtout en raison des positions différentes des ombres). De même, l'aspect de la végétation peut changer. Compte tenu de l'évolution du paysage, un certain nombre d'objets topographiques peuvent subir une évolution radiométrique importante (vieillessement des toitures).

2.4.4.3. *Considérations algorithmiques*

2.4.4.3.1. Paramétrage des méthodes

Comme dans tout processus d'analyse d'image, l'étude des paramètres doit être faite afin d'apprendre à maîtriser la méthode et à connaître ses limites. On visera notamment à réduire le nombre de paramètres critiques. Par exemple, on peut aisément remplacer un seuillage de coefficient de corrélation (qui est généralement sensible) par l'introduction de la corrélation croisée symétrique (voir points homologues dans les notions de base).

2.4.4.3.2. Nombre de points utiles par image

Des études ont montré qu'il n'y avait pas d'amélioration du résultat de l'aérotriangulation en prenant plus de 20 points par image (si les 20 points sont corrects). Dans le cas de la mesure automatique, il est nécessaire d'avoir plus de points pour pouvoir rejeter de façon fiable les points aberrants. Le nombre de points nécessaires est donc plus élevé.

2.4.4.3.3. Évaluation

L'évaluation de la qualité des points de liaison est basée sur la qualité de l'aérotriangulation issue de cette mise en place. Un premier critère de qualité se trouve être les résidus (en microns ou en pixels) des points de liaison. Une autre technique consiste à évaluer la qualité des points de liaison à l'aide de points terrain de référence. Ces points terrain peuvent être calculés à l'aide d'une aérotriangulation fiable ou de mesures terrain.

2.4.4.3.4. Combinatoire

Au moment de la recherche de points homologues entre deux images, on peut *a priori* tenter de mettre en correspondance un point quelconque de la première image, avec un point quelconque de la seconde image. Dans ce cas, la combinatoire du problème est importante. Dans le cas de n images, la combinatoire augmente encore. La combinatoire peut être fortement réduite en restreignant correctement l'espace de recherche des homologues d'un point donné.

Une technique classique pour cela est l'approche multi-échelle (déjà citée plus haut) : en effectuant préalablement les calculs à échelle réduite, on peut « prédire » la position d'un point à rechercher.

Une autre solution réside dans l'approche « géométrique »: si, pour une zone de recherche donnée, on ne considère comme candidats que les positions des points d'intérêt qui y sont présents, et pas tous les pixels de la région, alors la

combinatoire chute sensiblement. Toutefois cette méthode repose entièrement sur la répétabilité du détecteur de points d'intérêt.

2.4.5. Constitution automatique de tableau d'assemblage

L'objectif de la constitution d'un tableau d'assemblage est ici de calculer une information de position et d'orientation 2D des images de la mission, permettant de placer toutes ces images dans un même repère du plan, de telle sorte que les détails homologues de plusieurs images se trouvent à la même position dans ce repère commun. Pour deux images possédant un recouvrement, il est alors possible d'estimer un changement de repère permettant la prévision de la position d'un point homologue à un point p donné. Ainsi, le tableau d'assemblage, outre son utilité propre, sert de guide à la détection des points de liaison multiples.

2.4.5.1. Les images sont traitées après un fort sous-échantillonnage

Formellement, on considère un repère global au plan commun à toutes les images. C'est le repère de bloc. Ce que l'on recherche est, pour chaque image I , une similitude $s_I = R_I \circ T_I$ permettant le passage des coordonnées d'un point du plan (x_B, y_B) exprimées dans le repère de bloc, à sa position $p(x, y)$ dans l'image. Le problème est traité bande par bande, puis par un assemblage de bandes. Les techniques s'appuient sur la détection de points d'intérêt dans les images, et sur le calcul de l'orientation relative 2D d'un couple d'images dont la rotation relative est d'angle voisin de 0.

2.4.5.2. Détection de points d'intérêt

Cette partie de la méthode comporte deux paramètres : l'écart type de la gaussienne utilisée pour les lissages et dérivations, et la taille de la fenêtre de calcul associée à ces traitements.

2.4.5.3. Traitement d'un couple d'images d'orientation voisine

On suppose que les images possèdent une rotation relative d'angle voisin de 0, et qu'en raison du fort sous-échantillonnage, les disparités des points homologues de deux images sont toutes semblables. Deux points homologues peuvent donc être mis en correspondance sans problème de rotation. Pour deux points d'intérêt P_i de l'image I , et Q_j de l'image J on considère c_{ij} le coefficient de corrélation linéaire calculé sur une fenêtre carrée centrée sur chacun des points. Deux points P_i et Q_j sont jugés candidats homologues si pour P_i aucun autre point de J n'est plus ressemblant, et inversement. Formellement :

$$\begin{aligned} \forall j', c_{ij'} &\leq c_{ij} \\ \forall i', c_{i'j} &\leq c_{ij} \end{aligned} \quad [2.15]$$

On obtient ainsi des couples de points candidats pour les images I et J. Dans la pratique, beaucoup d'aberrations sont présentes. On opère donc un filtrage dans l'espace des disparités. Pour deux points candidats homologues, on définit le vecteur de disparité reliant ces deux points. Il s'agit du vecteur obtenu par différence des coordonnées des points. En accumulant ces vecteurs de disparité dans un histogramme, on obtient un nuage correspondant aux points de disparités similaires. Ces points constituent les points homologues retenus par la méthode.

Dans la pratique, le nuage est identifié par une fermeture morphologique avec un élément structurant carré 3 x 3. Après la fermeture, la composante connexe de poids le plus fort est identifiée, puis les points participant à cette composante sont retenus. À ce stade, on définit la performance d'un couple d'images : il s'agit du rapport entre le nombre de points participant au nuage de points et le nombre de points candidats homologues.

À partir de ces points, on estime aux moindres carrés la similitude permettant au mieux de relier les deux repères des images. Ceci permet de « plonger » la seconde image dans le repère de la première. Cette partie de la méthode comporte un seul paramètre concernant la corrélation : il s'agit de la taille de la fenêtre de calcul. La taille de l'élément structurant pour le filtrage est par ailleurs fixe.

2.4.5.4. Traitement d'une bande

On suppose connaître les N images constituant une bande. Le traitement précédent est appliqué aux N-1 couples successifs constituant une bande. On peut ainsi « plonger » toutes les images d'une bande dans le repère de la première image de cette bande. Toutes les bandes sont traitées indépendamment. Il n'y a aucun paramètre supplémentaire concernant cette étape.

2.4.5.5. Traitement d'un bloc

On suppose connaître l'ordre des bandes. On suppose également que les bandes possèdent un recouvrement latéral significatif (> 10 %). Cette partie de la méthode est réalisée itérativement : on colle la bande $n + 1$ au bloc constitué des n bandes précédentes. On considère donc que l'on dispose d'un bloc déjà constitué auquel on ajoute une bande supplémentaire. Ceci est réalisé en accrochant d'abord une image de la bande au bloc, puis en raccrochant les autres images de la bande, de proche en proche.

Pour raccrocher une image de la bande au bloc : pour une image I de la bande, on considère temporairement tous les couples entre I et les images J de la dernière bande du bloc. On considère également les couples entre I et les images J ayant subi une rotation de π . Pour tous les couples formés, on effectue le traitement de couples d'images étudié plus haut, puis on retient le couple possédant la meilleure performance. Si cette performance n'est pas suffisante, on répète l'opération jusqu'à trouver un couple assez performant. On obtient ainsi un couple en une image I_0 et J_0 permettant de « plonger » l'image I_0 dans le repère du bloc.

De proche en proche, on plonge les images de la bande dans le repère du couple. Pour contrer les accumulations d'erreurs, une optimisation du système est effectuée à intervalles réguliers (toutes les n_{opt} images), par une méthode décrite dans la suite.

Cette partie de la méthode introduit deux paramètres. Le premier est le seuil de performance au-delà duquel on accepte le couple d'image entre la bande et le bloc. Ce seuil est fixé à une valeur dont une étude précise a montré qu'il était tout à fait fiable. Le second est $n_{opt} = 5$. Cette valeur permet à la méthode d'être fiable tout en réduisant le nombre d'optimisations.

2.4.5.6. Optimisation de bloc

Cette optimisation a pour but d'affiner les positions et orientation des images d'un bloc. Ceci est nécessaire car il n'est pas possible de « coller » de façon rigide deux bandes en raison de déformations qui s'accumulent lors de la constitution de tableau d'assemblage intra-bande. On considère tous les couples d'images possédant un recouvrement. Pour chaque couple, les points homologues au sens précédent sont calculés. Une fonction de coût visant à réaliser l'objectif du tableau d'assemblage est minimisée. Techniquement cette fonction est la somme de fonctions de coût liées à chaque couple d'images. La fonction liée à un couple d'images vise elle-même à minimiser, pour chacun des points homologues du couple d'images, la distance des deux homologues dans le repère du tableau d'assemblage :

$$d(s_i^{-1}(P_i), s_j^{-1}(P_j))$$

Cette fonction est ajustée pour ne pas tenir compte des points aberrants, mais son comportement global est celui décrit ci-dessus. Il n'est pas utile d'explicitier les détails de cette fonction de coût.

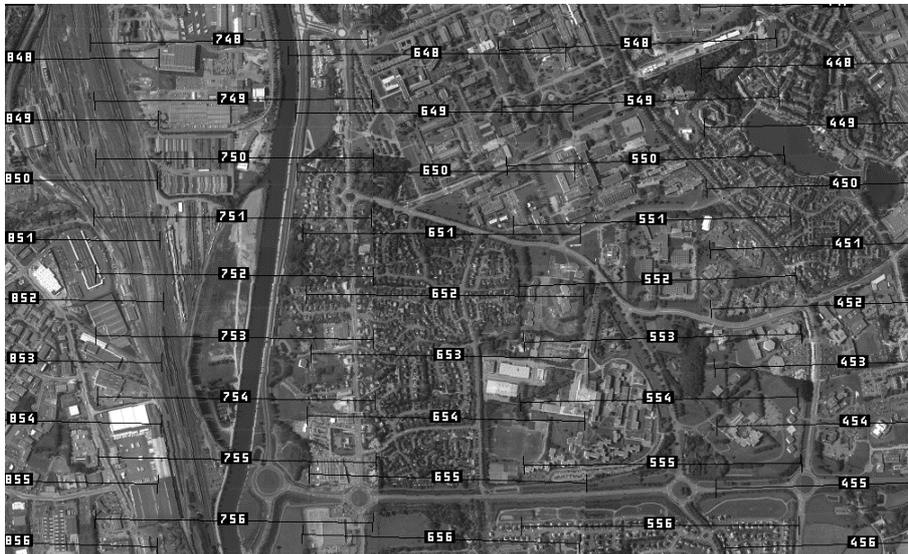


Figure 2.36. Extrait du tableau d'assemblage pour la mission Rennes. 14 bandes de 60 à 65 images ont été traitées. La précision géométrique des raccords de la mosaïque résultante est bonne : les raccords se font généralement avec une erreur de 1 ou 2 pixels

2.4.5.7. Étude des paramètres pour la constitution du tableau d'assemblage

Le comportement des paramètres est connu. Il s'avère notamment que pour certains d'entre eux, les valeurs sont fixes dans le sens où des chantiers significatifs ont été traités sans devoir les remettre en cause. Il s'agit de :

- la valeur de l'écart type de la gaussienne dans le calcul des points d'intérêt, ainsi que la taille de la fenêtre de calcul associée ;
- la taille de la fenêtre de corrélation pour le calcul des points candidats homologues ;
- la taille de l'élément structurant utilisé dans le filtrage dans l'espace des disparités.

Pour les autres paramètres, leur comportement est connu : leur variation permet dans tous les cas d'augmenter la fiabilité au prix d'un temps calcul plus élevé, ou inversement. Dans ce sens, les paramètres ne sont pas sensibles car leur influence sur les résultats est connue.

2.4.6. Conclusion

Le problème reste largement ouvert. Malgré cela, sous certaines conditions, des systèmes sont tout à fait capables de produire des points de liaison aptes à calculer des chantiers photogrammétriques. Ces outils peuvent être très utiles pour de gros chantiers.

Dans tous les cas, la photogrammétrie numérique aura besoin de tels outils, ce qui justifie amplement les développements actuels dans ce domaine : les zones problématiques (occlusions, fort relief, texture, changement de date, variation forte d'échelle...) restent des points difficiles pour ces méthodes.

2.5. Stations de travail photogrammétriques numériques⁶

2.5.1. Fonctions d'un appareil de restitution photogrammétrique

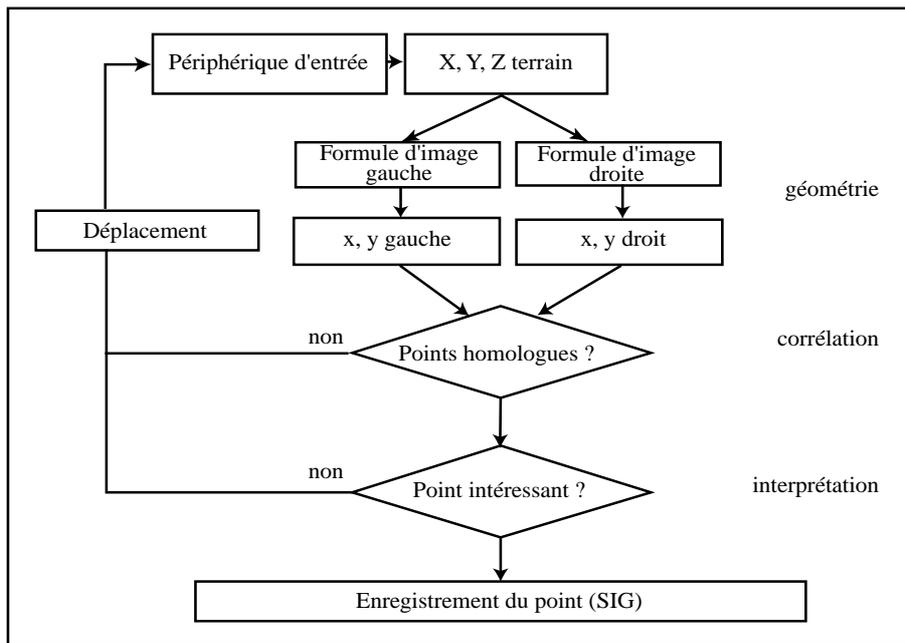


Figure 2.37. Principes de la restitution photogrammétrique

6. Cette section a été rédigée par Raphaëlle HENO et Yves EGELS.

Quel que soit son principe de réalisation, on retrouvera dans tout restituteur photogrammétrique un certain nombre de fonctions de base, que l'on peut représenter sur le schéma de fonctionnement suivant (cf. figure 2.37).

Ce principe de fonctionnement a été découvert dès les premières réalisations industrielles du début du XX^e siècle (restituteur de Von Orel chez Zeiss, Autographe de Wild). Dans les premiers systèmes, la fonction géométrique est réalisée par un calculateur analogique optique ou mécanique (d'où leur nom de restituteurs analogiques). Les deux autres fonctions font alors appel à un opérateur humain, dont la vision stéréoscopique assure la corrélation, et les connaissances culturelles l'interprétation.

Dans les années 1970, est apparue une première évolution sensible, le restituteur analytique : les formules d'image sont calculées analytiquement par un ordinateur, qui déplace les images devant un système optique grâce à des asservissements, car les images restant argentiques, la corrélation aussi bien que l'interprétation sont toujours confiées à un opérateur humain.

Aujourd'hui, les images sont devenues numériques, et peuvent être visualisées sur un simple écran d'ordinateur. De plus, les fonctions non géométriques sont de plus en plus souvent prises en compte par des moyens algorithmiques, et l'opérateur humain, donc la visualisation elle-même, deviennent alors inutiles. Dans ce dernier cas, il n'y a plus à proprement parler de restituteur photogrammétrique, mais uniquement des logiciels spécialisés fonctionnant sur des ordinateurs plus ou moins puissants.

Dans la suite, on se placera principalement dans le cas où un opérateur humain doit, sinon réaliser manuellement la totalité du processus photogrammétrique, tout au moins le superviser, en contrôler les résultats, et éventuellement reprendre manuellement certaines opérations.

2.5.2 Technologie

2.5.2.1. Matériel

Le passage au numérique en photogrammétrie a permis de s'affranchir des composants optiques et mécaniques de haute précision, qui maintenaient l'acquisition et l'entretien des systèmes à un niveau de prix élevé. Par ailleurs l'encombrement des machines est considérablement réduit.

Les systèmes de photogrammétrie numérique sont basés sur des stations de travail ou, de plus en plus, sur des PC. Les ordinateurs récents comportent en série la quasi-totalité des dispositifs nécessaires à la restitution photogrammétrique, et disposent aujourd'hui d'une puissance tout à fait

satisfaisante pour ces applications. Il ne faut pas oublier cependant que les données images manipulées ont généralement des tailles de plusieurs centaines de mégaoctets, et qu'un chantier nécessitera l'usage de dizaines, voire de centaines d'images. La capacité de stockage et les temps d'accès, devront donc être spécialement adaptés.

Seuls deux éléments indispensables ne sont pas encore aujourd'hui totalement standard, encore que le développement des jeux vidéo les rendent de plus en plus courants : le périphérique d'entrée et la visualisation stéréoscopique ; sur le plan de la conception informatique, la photogrammétrie numérique n'est guère plus qu'un jeu vidéo à finalité technique.

2.5.2.1.1. Périphérique d'entrée

Le système doit posséder un périphérique d'entrée de coordonnées en trois dimensions. Les solutions sont nombreuses, depuis les flèches du clavier, jusqu'aux manivelles et pédale des appareils photogrammétriques traditionnels, en passant par les contrôleurs d'immersion des stations de réalité virtuelle, les *joysticks*, ou les souris spécialisées. Seul le rapport prix/ergonomie peut les départager. La pratique établit que le mouvement doit être le plus continu possible, et que la commande en position est préférable aux commandes en vitesse (*joysticks*). Lorsqu'un périphérique spécialisé est développé, on en profite souvent pour lui adjoindre des boutons supplémentaires, permettant de simplifier l'emploi des commandes les plus fréquentes.

2.5.2.1.2. Visualisation stéréoscopique

Il doit également permettre la vision stéréoscopique, dans la mesure où l'on demandera à l'opérateur de réaliser des corrélations visuelles. Là aussi, plusieurs solutions existent, et certaines sont en cours de normalisation. Il s'agit de présenter à chaque œil l'image qui lui revient. Par ordre de confort croissant, on peut citer :

- l'affichage des deux images dans deux fenêtres séparées, que l'opérateur observera avec un système optique adéquat de type stéréoscope placé devant l'écran ;
- l'affichage simultané des deux images, l'une colorée en vert, l'autre en rouge, l'opérateur portant des lunettes complémentaires (procédé des anaglyphes) ;
- l'affichage alternatif et superposé des deux images, un dispositif à base de cristaux liquides ne laissant passer l'image que vers l'œil auquel elle est destinée.

Dans les systèmes professionnels, seule cette dernière solution est habituellement mise en œuvre, en raison de sa meilleure ergonomie. La fréquence de balayage de l'écran conditionne la qualité de la visualisation stéréoscopique, car l'affichage alternatif divise par deux la fréquence réellement perçue par l'utilisateur (une fréquence minimale de 120 Hz est nécessaire pour éviter la fatigue de l'opérateur). Elle est susceptible de plusieurs réalisations pratiques, suivant que les cristaux liquides sont placés directement devant l'œil, solution peu onéreuse, commandés par un fil ou par infrarouge, ou encore placés devant l'écran, l'opérateur portant juste une paire de lunettes polarisées passives (ce qui autorise notamment à regarder indifféremment l'écran stéréoscopique et l'écran de commande), solution la plus confortable mais aussi la plus coûteuse.

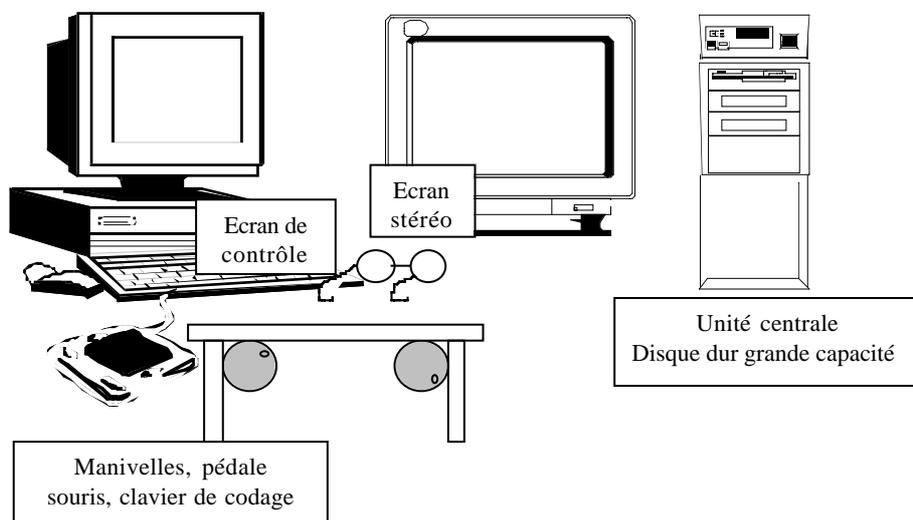


Figure 2.38. Schéma d'un système de photogrammétrie numérique

2.5.2.2. Algorithmes photogrammétriques

Il n'a pas fallu réinventer la photogrammétrie pour passer au numérique. Les équations posées sont les mêmes que sur les systèmes de photogrammétrie analytiques (équations de colinéarité ou de coplanéité).

De plus en plus de systèmes sont « multicateurs », c'est-à-dire qu'ils peuvent non seulement traiter les images aériennes, mais aussi des images de géométrie différente, comme les images issues de capteurs à balayage ou radar. La mise en place de ces images passe par des modèles mathématiques différents de la perspective photographique traditionnelle, éventuellement paramétrables par l'utilisateur.

Quelle que soit leur origine, une fois les images mises en place ou géoréférencées, la fonction « terrain→image » permet de leur transmettre en temps réel le déplacement terrain introduit grâce à la souris ou aux manivelles.

Par contre, contrairement aux restituteurs analogiques ou analytiques, la disponibilité de l'image sous forme numérique permet d'étendre dans des proportions considérables les possibilités d'automatisation de la photogrammétrie, nécessitant ainsi un rapprochement dans un même système des fonctions photogrammétriques et des fonctions de traitement d'image et de reconnaissance de formes.

2.5.3. Affichage des données image et vecteur

2.5.3.1. Affichage des images

Les images que l'on exploite ont toujours des tailles très supérieures à la dimension de l'écran (couramment 100 fois supérieures). Il faudra donc pouvoir déplacer commodément la fenêtre d'affichage dans l'image totale. Deux modes de déplacement sont possibles : soit l'image est fixe, et le curseur de mesure mobile ; soit le curseur est fixe et central, et l'image mobile. La mesure se fera en superposant un index de pointé, le ballonnet (dont la couleur pourra s'adapter automatiquement au fond d'image sur lequel il est placé).

2.5.3.1.1. L'image est fixe

La configuration image fixe/curseur mobile est la plus facile à implémenter : il suffit de déplacer quelques octets en mémoire vidéo. Mais l'ergonomie de cette solution est médiocre. Si les images ne sont pas retraitées géométriquement (rééchantillonnage épipolaire), la parallaxe transversale n'est éliminée qu'en un seul point, généralement le centre de l'image. Lors des déplacements du curseur, une parallaxe apparaît, ce qui rend le pointé sinon impossible, tout au moins imprécis. De plus, lorsque le curseur atteint le bord de l'écran, le système doit recharger une autre zone des images, ce qui interrompt le travail de restitution et déconcentre l'opérateur.

2.5.3.1.2. Le curseur est fixe

Plus gourmande en capacités de calcul, la configuration image mobile/curseur central est bien préférable. Si le déplacement des images est suffisamment fluide, et il doit l'être pour ne pas fatiguer l'opérateur, cette configuration rappelle les appareils analytiques, où les porte-clichés se déplacent de façon continue devant un ballonnet fixe. Les images sont généralement chargées en mémoire par tuiles, pour optimiser le temps des rechargements.

Le système propose parfois une vue globale (image sous-échantillonnée) pour les déplacements rapides dans le modèle.

En mode « exploitation du modèle », il est préférable de consacrer la surface maximum de l'écran stéréo à l'affichage des deux images. En revanche, lors des mesures de points d'aérotriangulation, il est pratique de pouvoir afficher pour chaque point toutes les images où il est présent (multifenêtrage).

2.5.3.2. Zooms et déplacement subpixellaire

Les zooms avant par interpolation bilinéaire ou au plus proche voisin servent à agrandir tout ou partie de l'écran courant. Leur but est habituellement de permettre des pointés d'une précision inférieure au pixel réel de l'image. Malheureusement, le zoom diminue simultanément le champ d'observation, et ajoute un flou important à l'image, ces deux effets dégradant sensiblement la qualité de l'interprétation.

Le déplacement subpixellaire de l'image permet de pallier simultanément ces deux inconvénients (mais cette fonction n'est généralement pas disponible dans les bibliothèques graphiques standards, ce qui nécessite une programmation spécifique).

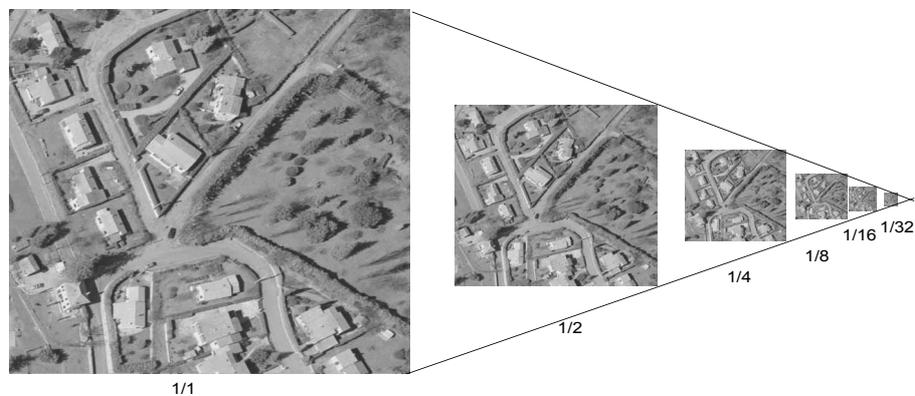


Figure 2.39. Pyramide de zooms arrière

Dans le cas où l'image est fixe, il suffit de rééchantillonner le curseur de façon à ce qu'il paraisse positionné entre deux pixels, ce qui nécessite qu'il soit lui-même formé de plusieurs pixels. Si l'image est mobile, c'est alors la totalité de l'image affichée qu'il faut recalculer.

S'agissant des zooms arrière, ils sont calculés à la volée ou générés à l'avance sous forme d'une pyramide d'images physiques de résolution décroissante (cf. figure 2.39). Dans le cas des pyramides d'images, on ne pourra pas dézoomer à n'importe quelle échelle.

2.5.3.3. Traitement d'images

Les boîtes à outils de base contiennent au moins des fonctionnalités de réglage du contraste, de la luminosité, d'inversion positif/négatif. Basés sur la table de couleurs (LUT) des images, ils fonctionnent en temps réel, et ne nécessitent pas la création de nouvelles images.

Il est parfois possible d'appliquer des filtres de convolution (rehaussement de contours, lissages...). Leur application étant assez gourmande en temps CPU, elle ralentit notablement le temps de chargement des images. De plus, ces filtres ont souvent des effets pervers sur le plan géométrique (déplacement de contours), dont l'influence sur la précision de la saisie peut être catastrophique. Pour plus d'efficacité, on préférera parfois calculer de nouvelles images.

2.5.3.4. Rééchantillonnage épipolaire

La visualisation d'un modèle stéréoscopique dont les images ont été acquises avec des angles sensiblement différents peut être fatigante pour les opérateurs, car les deux images ont des échelles et des orientations trop dissemblables. Les performances des techniques habituelles de corrélation sont d'ailleurs également dégradées par ce type d'images.

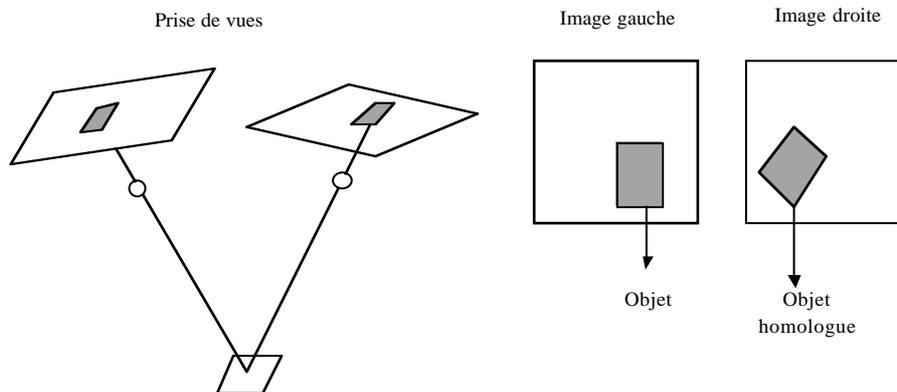


Figure 2.40. Visualisation de modèle stéréoscopique avec des rotations différentielles importantes

Dans ce cas, il est souhaitable de faire un rééchantillonnage épipolaire : on calcule l'homographie plane qui permet de passer de l'image initiale à celle qui aurait été obtenue si les axes des photos avaient été parallèles entre eux et perpendiculaires à la base (cas normal) ; après ce traitement, deux points d'une ligne de l'image gauche auront leurs homologues dans une même ligne de l'image droite, et la parallaxe transversale est constante. Les lignes épipolaires sont les intersections du faisceau de plans contenant la base et des deux plan image. Dans les images rééchantillonnées, ces droites sont parallèles, alors qu'elles étaient concourantes dans les images initiales.

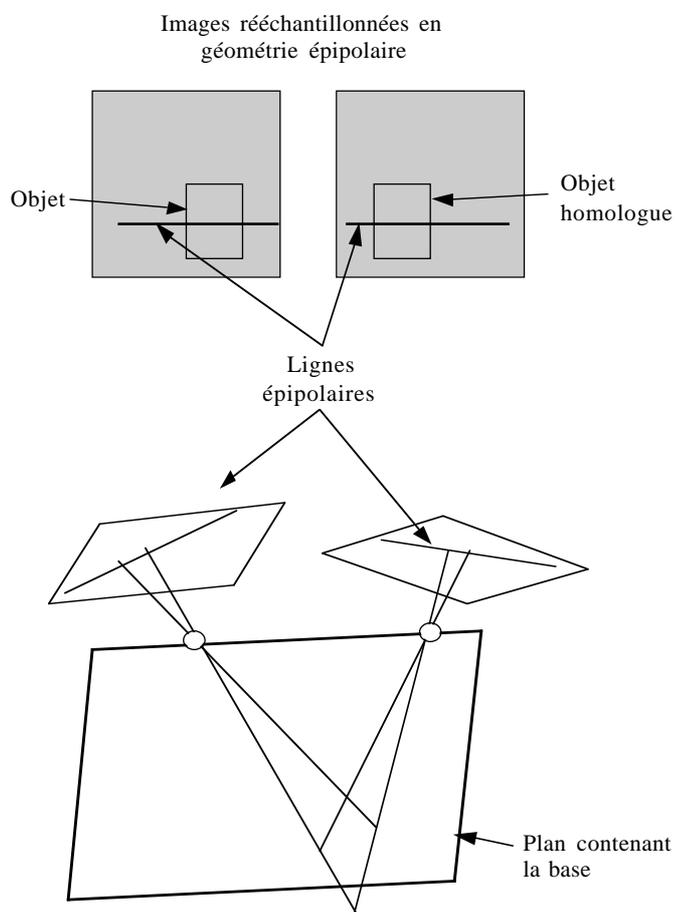


Figure 2.41. Rééchantillonnage épipolaire

Ce rééchantillonnage nécessite au minimum le calcul de l'orientation relative (et ne peut bien entendu pas être utilisé lors des mesures nécessaires à son établissement). Il génère deux nouvelles images dont il faut tenir compte dans la gestion de l'espace disque. Comme il n'est valable que pour un couple, il est possible de découper les images et de ne garder que les parties communes des images du modèle.

Le rééchantillonnage épipolaire améliore non seulement le confort de l'opérateur, mais il accélère aussi le travail des corrélateurs, puisque la zone de recherche des points homologues est réduite à un espace à une dimension (contre un espace à deux dimensions pour des images non rééchantillonnées). Certains systèmes de photogrammétrie numérique exigent de travailler avec des images en géométrie épipolaire pour l'extraction du MNT, d'autres effectuent ce rééchantillonnage à la volée, sans calculer d'images supplémentaires.

2.5.3.5. *Affichage des vecteurs*

Alors que sur les restituteurs analytiques cette fonctionnalité (très utile pour la mise à jour cartographique) était une option très onéreuse, nécessitant des adaptations optiques complexes, la superposition aux images des vecteurs en couleurs et en stéréo ne pose aucun problème sur les systèmes de photogrammétrie numérique.

2.5.4. *Fonctionnalités*

Les systèmes de restitution numérique offrent au minimum les mêmes fonctionnalités que les restituteurs analytiques. D'ailleurs, celles-ci peuvent pratiquement être réutilisées sans changement, excepté l'addition de fonctions d'aide au pointé, qui tirent profit de la nature numérique de l'image. Mais leurs caractéristiques permettent d'envisager de nombreuses extensions de leur emploi. On trouvera ici résumées les principales d'entre elles.

Contrairement à ce qui se passait dans les systèmes analogiques ou analytiques, la qualité géométrique des produits calculés sur le système de photogrammétrie numérique (base de données vecteurs, MNT, orthophotos) sera à peu près indépendante du système. Les algorithmes géométriques utilisés sont très proches, et plus aucune mécanique n'intervient. Il conviendra de veiller spécialement à la qualité des étapes suivantes :

- état des films ;
- qualité de la numérisation (type de scanner utilisé, maîtrise du paramétrage, pas de scannage) ;
- qualité des points d'appui ;
- qualité des mesures de points d'orientation relative et absolue.

2.5.4.1. *Gestion des données*

2.5.4.1.1. Images manipulées

Les systèmes de photogrammétrie numérique sont capables de travailler avec des images issues du scannage des photographies aériennes classiques (exemple : RMK TOP de Zeiss, RC de Wild...), de photographies terrestres scannées, ou bien directement avec des images issues de caméras numériques aériennes ou spatiales (exemple : la caméra numérique matricielle de l'IGN).

Les images noir et blanc 8 bits et couleur 24 bits sont couramment lues par les systèmes de photogrammétrie numérique. S'agissant des images comportant 12 bits ou plus par canal, qui sont significatifs pour les capteurs numériques, il arrive que seuls les 8 premiers bits de chaque pixel soient en fait utilisés. De toute façon, les écrans d'ordinateur courants n'affichent en réalité que 6 bits, ce qui est aussi la dynamique maximale réelle des films, et l'œil humain ne distingue qu'une partie des 256 niveaux de gris d'une image codée sur 8 bits, mais avec les capteurs numériques de grande dynamique, cette perte d'informations est dommageable pour les algorithmes de corrélation et de reconnaissance automatique des formes.

Il n'y a pas de consensus chez les fabricants de systèmes de photogrammétrie numérique sur le format informatique d'images à utiliser. Quasiment tous sont capables de lire et d'écrire du TIF (tuilé ou non), mais certains préconisent de le convertir au format propriétaire pour optimiser les temps de traitements. Il en est de même pour la compression JPEG : son utilisation directe ralentit certaines applications, d'autant plus que cette technique (dans son état actuel) empêche l'accès direct à des portions d'image. La compression par ondelettes donnant d'excellents résultats en terme de rapport « qualité de l'image compressée/réduction de volume », il est probable que les systèmes de photogrammétrie numériques l'adopteront rapidement.

2.5.4.1.2. Base de données photogramétrique

Les données nécessaires à l'exploitation de chantiers importants sont assez nombreuses et très volumineuses ; elles nécessitent une organisation rigoureuse, qui est prise en compte, soit par une hiérarchie prédéfinie de répertoires, soit par une base de données spécifique.

Les systèmes permettent généralement de stocker les données par projet, qui correspondent intuitivement à une emprise géographique donnée, à un thème, etc. Un projet peut alors correspondre à un répertoire de l'arborescence du système, dans lequel on trouve les fichiers nécessaires à la mise en place des modèles (fichiers de caméra, fichiers contenant des points d'appui, éventuellement des fichiers contenant des mesures image réalisées sur un autre

appareil...), ainsi que les fichiers générés par les calculs (paramètres de l'orientation interne, position et attitudes de la caméra pour chacun des clichés, matrices de rotation...). Les données sont gérées par modèle, ou image par image (un fichier par modèle, ou un fichier par image). Les images sont stockées dans ce même répertoire, ou pour plus de souplesse dans la gestion de la place disque, y sont simplement référencées.

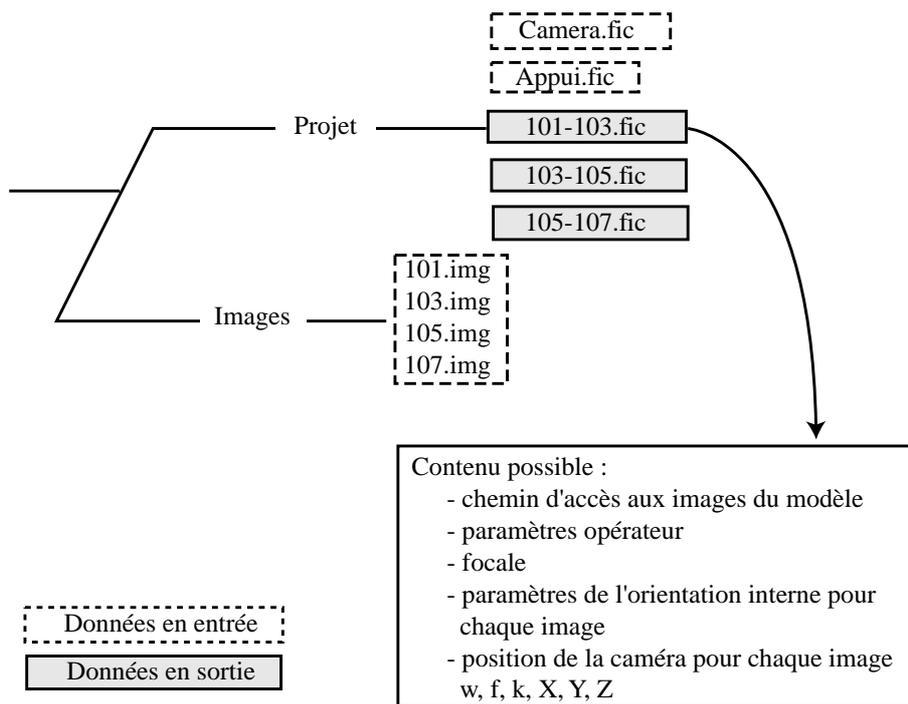


Figure 2.42. Exemple de base de données photogrammétrique

Les réseaux locaux (Ethernet 100, puis Giga Ethernet) permettent de travailler en architecture client-serveur : les images sont mises en place et stockées sur une machine « serveur », tandis que les machines « clientes » les utilisent (visualisation, traitements, saisie) *via* le réseau, ce qui demande des mécanismes de contrôle de la cohérence des données.

2.5.4.2. Fonctionnalités photogrammétriques de base

2.5.4.2.1. Mise en place d'un couple stéréoscopique

Orientation interne

La phase d'orientation interne est toujours nécessaire pour les images issues de la numérisation des photographies aériennes classiques. La connaissance de la position théorique des repères de fond de chambre (*via* le certificat d'étalonnage de la caméra) et de la forme de ces repères a permis d'automatiser au moins partiellement cette phase de mesures, grâce à la corrélation.

Il est parfois demandé à l'opérateur de rechercher manuellement les premiers repères, le système reconnaissant les suivants automatiquement. Les pointés interactifs sont assistés à la demande par corrélation automatique.

L'orientation interne est complètement automatique sur certains systèmes.

Orientation externe

La plupart des systèmes calculent simultanément les orientations relative et absolue, par l'utilisation de l'équation de colinéarité. Même si la mesure des points d'orientation relative est assistée par corrélation, voire complètement automatique dans les cas où l'on utilise une répartition théorique des points, il est toujours impératif qu'un opérateur repère manuellement la position des points d'appui.

L'élimination de la parallaxe en X et en Y se fait généralement en bloquant l'une des deux images, et en déplaçant l'autre avec le système de positionnement (manivelles, souris...). Une fois le choix de l'image fixe fait, on peut demander un pointé par corrélation. Dans certains cas, celui-ci peut échouer (couvert végétal, textures périodiques, zones homogènes), l'assistance de la corrélation ne dispense pas d'un pointé de l'opérateur, qui, s'il est un peu moins précis, reste malgré tout beaucoup plus fiable.

Dès que suffisamment de points sont mesurés (6 points pour avoir des résidus d'orientation relative, 3 points terrain), on peut lancer un premier calcul. La fonction « terrain → image » est alors connue, et permet de se déplacer automatiquement sur un point d'appui supplémentaire par exemple. Lorsque les mesures sont suffisamment surabondantes, l'examen des résidus de calcul permet de contrôler la qualité de la mise en place et de la valider.

2.5.4.2.2. Mesures et saisie dans un couple stéréoscopique

La saisie est bien souvent limitée à un couple stéréoscopique ; le chargement automatique des images du couple voisin en bord d'images est

parfois possible, mais nécessite bien sûr d'avoir toutes les images concernées chargées sur l'appareil.

Les fonctionnalités de base du système de photogrammétrie numérique sont :

- saisie de points, lignes, surfaces ;
- destruction, édition de ces structures.

Idéalement, le système de photogrammétrie numérique est interfacé avec un SIG qui gère les données saisies dans une base de données.

Toutes les fonctionnalités des SIG (analyses spatiales, requêtes, cartographie multi-échelle) sont alors directement utilisables après la saisie des données. Les données saisies sont structurées selon une topologie plus ou moins élaborée (trait, spaghetti, partage de géométrie, structure objet).

Les données extraites des restituteurs numériques sont systématiquement tridimensionnelles. Mais rares sont encore les SIG qui permettent de gérer complètement une topologie « vrai 3D » (gestion des faces verticales, des surplombs, des franchissements), et non pas seulement du 2,5D (topologie planaire, un seul attribut Z par point).

Les SIG proposent généralement des interfaces vers les formats les plus courants (DXF, DGN, Edigéo...).

2.5.4.3. Fonctionnalités photogrammétriques évoluées

2.5.4.3.1. Aérotriangulation

L'aérotriangulation permet de géoréférencer simultanément toutes les images d'un chantier, en tirant parti des recouvrements qu'elles ont en commun, avec un nombre minimum de points d'appui. Cette opération commence par une phase de mesure des coordonnées image d'un certain nombre de points vus sur le plus grand nombre possible d'images. Ensuite un calcul en bloc permet de déterminer l'ensemble des paramètres photogrammétriques du chantier. Les modules de calcul d'aérotriangulation disponibles sur les systèmes de photogrammétrie numérique utilisent les mêmes formules que l'aérotriangulation analytique et sont la plupart du temps identiques.

La mesure des points d'appui, toujours interactive, est facilitée par le multifenêtrage : une fois un point mesuré sur une image, le système peut afficher dans des mini-fenêtres toutes les images susceptibles de le contenir, zoomées directement sur la zone concernée. Ceci est rendu possible par la connaissance *a priori* des recouvrements entre images et de leur position les unes par rapport aux autres, déterminée grâce au GPS, ou *via* un tableau

d'assemblage défini par l'opérateur. Celui-ci n'a plus qu'à mesurer le point dans les fenêtres où il est présent, en mono, stéréo ou en demandant une assistance par corrélation.

La mesure des points de liaison est en revanche très automatisée : on peut remplacer la mesure précise d'un opérateur humain par des mesures par corrélation automatique en très grand nombre, soit par un choix de l'opérateur, aidé du multifenêtrage, soit par sélection automatique de points de liaison par un algorithme adéquat. Après filtrage des points faux, il reste normalement suffisamment de points pour assurer la tenue du bloc. Cette application est détaillée dans le chapitre 3

2.5.4.3.2. Altimétrie et modèle numériques de terrain

Les MNT peuvent intervenir soit comme donnée d'aide à la restitution, soit comme produit lui même. De plus, les stations photogrammétriques numériques sont parfaitement adaptées au contrôle et à la correction de MNT obtenus par des méthodes automatiques.

Dans le cas où un MNT préexiste (résultat d'une restitution antérieure, d'une conversion de courbes de niveau...), il permet d'asservir la marque de mesure au MNT, ce qui libère l'opérateur d'une tâche fastidieuse. Cependant, cette aide est assez limitée, ne serait-ce qu'en raison de la généralisation fréquemment excessive des modèles disponibles, et de la non-représentation du sursol.

De façon intermédiaire, il est possible d'aider le restituteur grâce à un corrélateur en temps réel, qui remplace le pointé altimétrique. On utilise dans ce cas un corrélateur fonctionnant dans l'espace objet, asservissant l'altitude au maximum de corrélation détecté sur la verticale déterminée par le système de commande.

La confection complète d'un MNT est habituellement laissée à un processus non interactif, éventuellement guidé par quelques mesures manuelles initiales ; en effet, la saisie manuelle d'un MNT est une opération longue et très fastidieuse. On préfère limiter les interventions humaines au contrôle et à la correction. Le contrôle se fait souvent visuellement, essentiellement par examen stéréoscopique des courbes de niveau calculées à partir du MNT superposées au couple d'images. Pour les corrections, chaque concepteur a imaginé des solutions différentes, allant du recalcul local avec des paramètres différents, avec intégration éventuelle de pointés humains complémentaires, jusqu'à des opérations beaucoup plus brutales sur les altitudes elles-mêmes : resaisie, arasement, creusement, interpolation.

2.5.4.3.3. Aides à l'interprétation

Les fonctionnalités d'extraction automatique ou au moins semi-automatique de données sont attendues avec impatience par les utilisateurs, qui ont investi dans les systèmes de photogrammétrie numérique en espérant que les algorithmes des chercheurs permettraient rapidement d'améliorer leur productivité. Hormis l'extraction automatique des MNT, et le rendu cartographique de courbes de niveau, aucun outil n'est encore industriellement implémenté sur les systèmes de photogrammétrie numérique. Le chapitre 3 évoquera l'état de la recherche dans ce domaine.

2.5.4.3.4. Orthophotographies, vues perspectives...

Les stations photogrammétriques numériques sont souvent vues comme des machines à fabriquer des orthophotos. En réalité, hormis pour les phases préparatoires (points d'appuis de l'aérotriangulation, contrôle et la correction des MNT) vues au paragraphe précédent, il est beaucoup plus économique de mettre en œuvre un logiciel spécialisé fonctionnant sur un ordinateur standard.

2.5.4.4. *Limitations*

Les stations de travail photogrammétriques permettent, fondamentalement, de faire exactement le même travail que les appareils de restitution analytiques qu'ils vont progressivement remplacer. Néanmoins il est certain que le confort oculaire des opérateurs n'est pas, et de loin, équivalent à celui de leurs prédécesseurs. Mais comme ces matériels suivent directement les possibilités des PC pour les jeux vidéo, où le problème est assez voisin, on peut s'attendre à des améliorations du confort de vision, point que les employeurs sont de plus en plus obligés de prendre en compte.

En contrepartie, ces stations permettent par ailleurs des traitements jusqu'ici impossibles à faire sur le même poste de travail, comme la réalisation d'orthophotographies.

Notons enfin que ces stations exigent évidemment des images sous forme numérique, qui pour le moment sont surtout fournies par numérisation de clichés argentiques, ce qui implique une phase de travail de digitalisation actuellement assez coûteuse, et un délai supplémentaire important. La généralisation de l'emploi de caméras digitales fournissant des images aériennes directement numériques est de nature à faire évoluer cette situation. Dans la situation intermédiaire qui prévaut à la date de rédaction de cet ouvrage, c'est essentiellement le prix d'acquisition et d'entretien des matériels (qui sont de simples PC à peine plus puissants que les appareils de base) et des logiciels (qui se démocratisent considérablement) qui rend ces stations attractives.