

# "Redressement de façades", proposition et étude d'une nouvelle méthodologie pour l'aide à la reconstruction automatique des bâtiments

VESTRI Christophe □◇

DEVERNAY Frédéric □

*ISTAR* □  
*Les espaces de Sophia F*  
*80 route des Lucioles*  
*B.P. 282*  
*06905 Sophia-Antipolis Cedex*  
*FRANCE*  
*fax: +33 (0) 4 93 95 83 29*  
  
*vestri@istar.fr*  
*devernay@istar.fr*

*INRIA-ROBOTVIS* ◇  
*2004 routes des Lucioles*  
*B.P. 93*  
  
*06902 Sophia-Antipolis Cedex*  
*FRANCE*  
*tél: +33 (0) 4 93 65 78 56*  
*fax: +33 (0) 4 93 65 78 45*  
*Christophe.Vestri@sophia.inria.fr*

**Mots clés :** redressement de façades, reconstruction de bâtiments, imagerie aérienne, imagerie spatiale, stéréoscopie, corrélation, reconstruction 3D, extraction de bâtiments.

**Résumé :** Nous présentons dans cet article une nouvelle méthodologie qui nous permet de prendre en compte les informations introduites par les façades de bâtiments. Ces surfaces correspondent à des zones très fortement distordues des images et sont difficiles à mettre en correspondance entre les images. Nous proposons ici l'application d'une transformation, le "redressement des façades", à chacune des images initiales. Dans chacune des nouvelles images construites, nous pouvons mettre en correspondance les façades par un algorithme de corrélation et reconstruire ensuite le bâtiment complet. Nous présentons les fondements de cette nouvelle méthodologie permettant de construire des images dans lesquelles les façades sont toutes perçues de manière identique. La seule hypothèse réalisée sur la nature de ces surfaces est leur verticalité. Nous présentons tout d'abord l'approche bidimensionnelle (dans une courbe épipolaire) du problème puis son extension tridimensionnelle. Tout au long de l'étude, les images reconstruites, ou "redressées", sont présentées. Nous finissons l'étude du concept présenté en montrant quelques résultats de reconstructions.

# 1 Introduction

Cet article présente un nouveau concept pouvant être utilisé dans un processus de reconstruction automatique de bâtiments. La méthode que nous présentons permet l'extraction des limites ou des bords des bâtiments. L'idée directrice est d'utiliser les informations relatives aux projetés des façades dans les différentes images. Ces parties des images correspondent à une quantité importante d'informations qui sont généralement négligées car la géométrie du système les rend difficilement utilisables. L'approche que nous présentons dans cet article va permettre d'obtenir de nouvelles images de la même scène dans une géométrie différente qui sera adaptée pour l'utilisation de ces façades. La seule hypothèse réalisée sur la nature des surfaces 3D recherchées (nos façades) est leur verticalité. La mise en correspondance des façades pour retrouver le plan 3D vertical du bâtiment étudié (et donc le bord du bâtiment) pourra dans ces nouvelles images être réalisée par un algorithme de corrélation classique.

La section suivante présente rapidement les principales approches concernant la reconstruction des bâtiments et des villes. Nous rappelons ensuite les principes de la mise en correspondance par corrélation et mettons en évidence les difficultés que rencontre ce type d'algorithme dans le monde de l'imagerie aérienne. Nous verrons dans la section 4 la présentation géométrique du problème que nous souhaitons résoudre. La section suivante présente l'étude du problème le long d'une courbe épipolaire, c'est le problème bidimensionnel. La section 6 s'intéresse au problème général avec des images initiales obtenues à partir d'un capteur ou modèle de projection générique. Nous verrons ensuite des résultats de reconstruction obtenus.

## 2 Extraction et reconstruction de bâtiments

L'extraction, puis la reconstruction de bâtiments à partir d'imagerie optique aérienne ou spatiale est un problème qui est traité depuis déjà de nombreuses années. La nature de la scène observée (discontinuités et irrégularité de la surface, occultations ou présence d'ombres, etc...) rend ce problème difficile et aucune solution ne peut le résoudre de manière directe, automatique et complète. Les approches existantes résultent généralement de la combinaison de différentes méthodes d'analyse d'image, de mise en correspondance ou de modélisation de données [2] [3] [4] [5] [9] [10] [12] [14] [15] [16] [17] [18].

L'une des étapes primordiale est celle de la stéréoscopie ou de la mise en correspondance des points images. Elle permet de calculer la position du point matériel dans l'espace 3D qu'ils représentent. On distingue gé-

néralement trois classes de méthodes: les méthodes à base de primitive ("feature-based"), les méthodes surfaciques ("area-based") et les méthodes énergétiques. On peut trouver une bonne revue de ces méthodes dans [7]. Nous verrons dans la suite que notre approche permettra grâce à la structure particulière des bâtiments d'utiliser des méthodes de type surfacique ou basées sur des similarités de voisinages pour retrouver la position 3D des façades et donc les bords des bâtiments. Nous présentons dans ce qui suit la méthode dite de corrélation que nous avons choisie.

## 3 La mise en correspondance des bâtiments par corrélation

La majeure partie des processus de mise en correspondance d'images aériennes ou spatiales est basée sur le calcul d'un critère de corrélation le long des lignes des images rectifiées (ou épipolaires, on peut se référer à [8] ou [13] pour la géométrie des capteurs ou d'un système de plusieurs caméras). Ces méthodes admettent des modèles de luminance et de réflexion simples. Elles permettent de retrouver principalement les surfaces dites "fronto-parallèles" de la scène. L'utilisation de photographies aériennes nous entraîne dans une configuration particulière du système de caméras. Les prises de vues sont réalisées le long de lignes de vol effectuées à une altitude importante. Les surfaces de notre scène fronto-parallèles au plan image du capteur correspondent dans ce cas à celles qui possèdent un caractère horizontal, autrement dit les toits et le sol.

### 3.1 Principes de base de la corrélation

Une des méthodes classiques de mise en correspondance de points est l'utilisation d'un critère de corrélation. C'est une méthode de type surfacique. Ces méthodes sont nées de la constatation que deux images sont généralement *localement* semblables autour des points se correspondant. La projection de la surface 3-D associée aux voisinages de ces points est similaire, d'où la dénomination de méthodes *surfaciques*. Le principe de ces méthodes est généralement le suivant: (1) on se fixe un point d'intérêt dans l'une des images, (2) à l'aide d'un critère de ressemblance à maximiser, on recherche le(s) point(s) homologue(s) lui correspondant dans l(es) autre(s) image(s) de telle manière que les contraintes imposées soient vérifiées. On peut se référer à [6] pour une présentation complète de l'algorithme.

Un certain nombre d'hypothèses ou de contraintes sont utilisées pour retrouver les voisinages qui se correspondent (contrainte lambertienne, surfaces fronto-parallèles et continues, continuité figurale, etc...). Dans la pratique, ces contraintes sont plutôt lâches car elles dépendent principalement de la nature de la scène et

ne sont pas réalistes. On ne peut donc les satisfaire pleinement mais on s'en sert pour obtenir une solution plus ou moins cohérente avec la réalité.

Les méthodes surfaciques ont l'avantage sur les méthodes à base de primitives de fournir une représentation dense de la scène 3D observée. Ces méthodes se proposent généralement de fournir une solution pour chaque point d'intérêt. On note toutefois que ces méthodes sont mises à défaut par les discontinuités de profondeur qui sont très présentes dans le cas d'une ville.

### 3.2 Problèmes liés aux façades en stéréoscopie aérienne

La mise en correspondance par corrélation dans le cadre de notre application a déjà été traitée dans de nombreuses publications dont [2] [5] [9]. Les principales difficultés apparaissent aux bords des bâtiments. Elles sont dues à la présence de discontinuités de surfaces. Les façades des bâtiments visibles dans les images vont engendrer deux autres catégories de problèmes. Tout d'abord, elles sont invisibles dans la région centrale des images puisqu'elles sont perpendiculaires à l'axe optique. Elles ne peuvent dès lors nous apporter aucune information supplémentaire au processus de mise en correspondance ou de reconstruction. Lorsque nous nous éloignons du centre optique de l'image, les façades apparaissent progressivement et leurs surfaces deviennent comparables à celles des toits dans le bord des images. Elles peuvent alors nous apporter une information complémentaire ou supplémentaire.

Le problème auquel nous nous sommes confrontés est l'utilisation correcte des images de ces façades. Les images des façades de deux bâtiments d'altitude identique ne possèdent pas la même taille dans l'image. Cette taille dépend de la position de celle-ci par rapport au centre optique. On observe ainsi un effet de dilatation ou de distorsion des images de façades plus important lorsqu'on s'éloigne du centre optique. Lorsque nous utilisons deux prises de vues distinctes d'une même scène, les images d'une même façade de bâtiment ne sont pas à la même distance des centres optiques des images. Les différentes images correspondant à la façade sont alors de tailles différentes.

Lors de la mise en correspondance de deux prises de vue distinctes, les surfaces visibles d'une même façade sont donc fortement distordues l'une vis-à-vis de l'autre dans les images (cf figure 1). Elles sont très difficiles, voire impossibles, à apparier. Il conviendrait alors de transformer les images des façades de telle manière que la surface soit perçue de manière identique dans chacune des vues. Nous allons voir dans la suite qu'à partir de l'hypothèse de verticalité des façades, un échantillonnage particulier du plan image nous permet d'éviter les

problèmes que nous venons de voir. Nous verrons par la suite qu'il sera possible d'obtenir de nouvelles images de la scène dans lesquelles nous pourrions apparier les façades et retrouver ainsi les points 3D de la scène qui leur correspondent.

## 4 Présentation géométrique du problème posé

Nous allons considérer le problème à deux dimensions suivant : on se place tout d'abord dans le plan passant par le centre optique et une ligne d'un capteur. On suppose ensuite utiliser un modèle de capteur simple qui est celui du modèle de projection perspectif faible ("scaled orthographic projection", cf [1] pour une étude détaillée sur l'approximation des modèles de projection). La particularité de ce modèle de projection est de supposer la scène observée comme étant à une distance fixe. Cette simplification réalisée sur le modèle de capteur nous apporte des propriétés géométriques qui vont nous permettre de comprendre le problème que nous voulons résoudre. Nous le présentons dans la figure 2.

La scène observée par le capteur est donc ici considérée comme étant une droite parallèle à la ligne image. La distance séparant la scène du centre optique est fixe et notée  $Z_c$ . La focale du capteur est notée  $f$ . Dans la suite et conformément à la figure 2,  $x$  représente l'abscisse image,  $X$  l'abscisse du point 2D dans la scène et  $\alpha$  l'angle défini par le rayon optique du capteur et la droite de projection correspondant au point étudié.

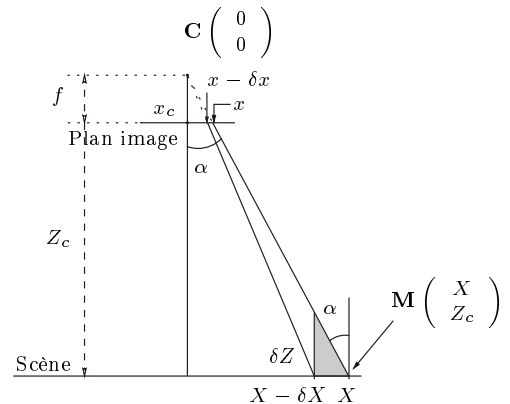


FIG. 2: Etude du modèle de projection en perspective faible.

La scène et le plan image étant parallèles, on déduit que les éléments  $\delta x$  et  $\delta X$  sont proportionnels, ou que le rapport de ces deux éléments est constant en tout point (nous appelons ici un "élément" une variation



FIG. 1: La paire d'images présente les problèmes que nous rencontrons avec les façades. Les surfaces verticales sont déformées et n'ont pas la même taille dans les différentes vues. Leur appariement est difficile.

(ou un déplacement) le long d'un axe ou d'une direction donnée) :

$$\frac{\delta X}{\delta x} = \frac{Z_c}{f} = \text{constante} \quad (1)$$

L'axe optique étant vertical, les surfaces verticales (correspondant à nos façades) sont toutes parallèles à celui-ci. L'angle  $\alpha$  que nous avons défini précédemment correspond aussi à l'angle formé entre la droite de projection ( $CM$ ) et la droite verticale passant par le point  $M$  de la scène (cf figure 2). Nous pouvons alors écrire :

$$\tan \alpha = \frac{x - x_c}{f} = \frac{X - X_c}{Z_c} = \frac{\delta X}{\delta Z} \quad (2)$$

Ces deux équations vont nous permettre d'expliquer et de comprendre le problème posé. La première équation (1) nous permet d'affirmer que dans cette configuration, les variations  $\delta x$  et  $\delta X$  sont linéairement dépendantes. Lorsque l'image est échantillonnée de manière régulière (à  $\delta x = \text{constante}$ ), on peut affirmer qu'à un élément unitaire  $\delta x$  (1 pixel) de l'image correspond parfaitement (ou proportionnellement) un élément de surface horizontale de la scène (à la distance fixe  $Z_c$ ).

La seconde équation (2) permet de définir la relation liant un élément horizontal ( $\delta X$ ) à l'élément vertical lui correspondant ( $\delta Z$ ) et réciproquement. Le lien entre ces deux grandeurs dépend par contre de la position du point  $M$  de la scène par rapport à l'axe optique (ou de la distance par rapport au point principal de l'image).

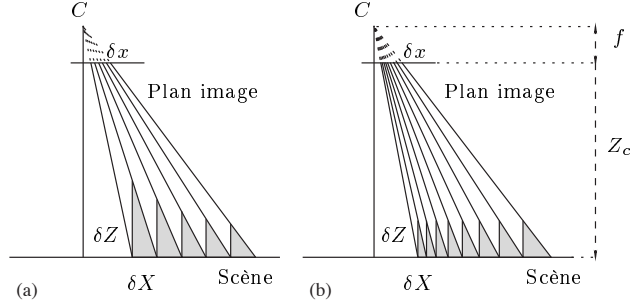


FIG. 3: Choix du critère d'échantillonnage : (a) échantillonnage avec  $\delta X$  constant, (b) échantillonnage avec  $\delta Z$  constant. Dans le premier cas, la taille des éléments verticaux perçus diminue en s'éloignant du centre optique. Dans le second cas, la taille des éléments horizontaux perçus augmente en s'éloignant du centre optique.

Partant de là, deux cas distincts se présentent : (1) on veut préserver (ou privilégier) les éléments horizontaux, (2) on veut préserver les éléments verticaux. Le premier cas (1) est présenté dans la figure 3 (a). On veut que l'image preserve les éléments horizontaux, c'est-à-dire que chaque variation élémentaire dans l'image (de 1 pixel, ou d'un élément  $\delta x$ ) corresponde à un élément horizontal **fixe** (ou constant) de la scène. En imposant la contrainte  $\delta X = K$ , on déduit de manière triviale que l'image doit être échantillonnée elle aussi de manière constante ( $\delta x = \frac{f\delta X}{Z_c} = \frac{fK}{Z_c}$ ). Ce premier cas correspond à l'échantillonnage couramment réalisé du plan image (ou d'une ligne dans le cas présent) pour obtenir les différentes vues de la scène. Dans ce cas d'échantillonnage, les surfaces fronto-parallèles à l'image qui sont échantillonnées de manière constante sont les toits et le sol de la scène. Il est alors possible de mettre en correspondance par corrélation les surfaces horizontales. Si l'on considère cette fois les éléments verticaux correspondant au même pixel de l'image, ils seront, quant à eux, représentés avec des tailles irrégulières ( $\delta Z = \frac{fK}{(x-x_c)}$ ). Plus on est proche du centre optique et plus l'aire de surface verticale représentée par un pixel est grand. D'un pixel à l'autre, la surface verticale observée est complètement différente. Il est donc impossible de mettre en correspondance les façades dans la configuration présente.

Dans le second cas (2), on décide de préserver les éléments verticaux (cf figure 3 (b)). On souhaite alors imposer  $\delta Z = K$  (où  $K$  est une constante choisie). Sous cette contrainte, on peut calculer les tailles des éléments horizontaux  $\delta X$  et  $\delta x$  correspondant à chaque élément vertical :  $\delta X = \frac{K(x-x_c)}{f}$  et  $\delta x = \frac{K(x-x_c)}{Z}$  qui dépendent l'un et l'autre de leur position par rapport à l'axe optique. Pour discrétiser (ou échantillonner) l'image de manière à ce que chaque élément  $\delta x$  représente un élément vertical de la scène unitaire (ou constant), il faut utiliser une grille irrégulière ou un pas d'échantillonnage variable comme le montre la figure 3 (b).

Dans un contexte plus général on peut à présent considérer des éléments d'orientation quelconque. Nous cherchons dans ce cas à échantillonner l'image de manière à obtenir des pixels images représentant des éléments de taille constante d'une surface d'inclinaison fixe par rapport à l'axe optique. Cette inclinaison est définie par rapport à la scène par l'angle  $\theta$  comme le présente la figure 4. L'équation 2 peut être réécrite comme suit :

$$\tan \alpha = \frac{x - x_c}{f} = \frac{\delta X - \delta S \cos \theta}{\delta S \sin \theta} \quad (3)$$

De la même manière que précédemment, nous choisissons de rééchantillonner l'image en imposant la contrainte  $\delta S = K$  (où  $K$  est une constante choisie). On obtient alors  $\delta X = \frac{K}{f}((x - x_c) \sin \theta + f \cos \theta)$  et  $\delta x =$

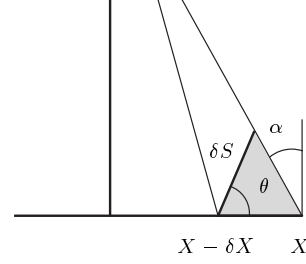


FIG. 4: *Etude des surface de la scène dont l'inclinaison par rapport à l'axe optique est constante.*

$\frac{K}{Z_c}((x - x_c) \sin \theta + f \cos \theta)$ . Une fois encore, pour que chaque pixel de l'image corresponde à un élément de surface (non fronto-parallèle) de la scène fixe, il faut que le pas d'échantillonnage tienne compte de la position de la surface par rapport au centre optique.

Partant de ce que nous venons de voir, le problème que nous cherchons à résoudre est d'obtenir des images préservant les surfaces verticales pour nous permettre de les mettre en correspondance par corrélation. Les images dont nous disposons sont par contre généralement échantillonnées à pas constant. Nous chercherons dans la suite la transformation qui nous permettra de passer d'une image échantillonnée à pas constant (ou préservant les surfaces horizontales comme dans la figure 3 (a)) à une image préservant les surfaces verticales (comme dans la figure 3 (b)).

## 5 Résolution du problème bidimensionnel

### 5.1 La méthode générale

Le but est de construire des images dans lesquelles nous allons pouvoir appliquer un algorithme de mise en correspondance par corrélation pour appairer les façades. Nous avons vu dans la section précédente que cela sera rendu possible avec des images "préservant" les éléments verticaux. On souhaite maintenant passer d'une image standard (échantillonnée à pas constant) à l'image souhaitée. L'idée est de déformer cette image pour que chaque pixel représente un élément vertical de taille constante pour toutes les images. Nous appellerons dans la suite l'idée présentée ci-dessus : *redressement des façades*.

### 5.2 Mise en équation

Le cas bidimensionnel présenté précédemment peut représenter plusieurs types de sous-problèmes liés à des modèles de capteurs précédents. (1) On peut considérer ce problème comme l'étude le long d'une droite épipo-

laire avec le modèle de capteur 3D sténopé. (2) On peut aussi considérer que l'on considère une acquisition du capteur 2D linéaire (ou pushbroom, cf [11]). Enfin, (3) on peut présenter ce problème comme étant l'étude du phénomène le long d'une courbe épipolaire de n'importe quel capteur ou modèle de projection. La caractéristique principale de tous les capteurs standards ou modèles de projection est d'échantillonner l'image à pas constant. Ce sont les raisons qui nous amènent à traiter le cas bidimensionnel dans cette section.

**Hypothèse 1** *Toutes les façades des bâtiments présents dans la scène correspondent à des surfaces verticales.*

Cette première hypothèse sur la structure de notre scène est très importante puisqu'elle va nous permettre à la fois de justifier et de guider notre démarche. Sous cette hypothèse, la transformation que l'on cherche peut s'identifier comme étant la fonction qui permet de passer d'une image échantillonnée à un pas constant à une image échantillonnée de telle sorte que chaque pas (ou pixel) de l'image représente un élément vertical fixe comme nous avons pu le voir. Cette transformation va permettre le passage de l'image originale (échantillonnée à  $\delta x$  constant comme dans la figure 3 (a)) à l'image souhaitée (échantillonnée à  $\delta z$  constant comme dans la figure 3 (b)).

La recherche des points homologues entre les différentes vues est réalisée le long des lignes épipolaires que nous considérons. On suppose dans cette section que chacune de ces lignes est indépendante dans le sens où nous traitons chaque ligne indépendamment de ses voisines. Ceci permet de faire abstraction de la seconde composante des "images épipolaires" utilisées (les colonnes). La transformation cherchée est restreinte elle aussi le long des lignes. On se retrouve, dans cette géométrie, à rechercher une transformation à une variable.

Considérons les images  $I_1$  et  $I_2$  comme étant deux lignes images correspondant respectivement aux deux prises de vues *Photo1* et *Photo2* distinctes d'une même scène  $S$ . On recherche les deux fonctions  $G_i$ , pour  $i = 1, 2$ , appliquées respectivement aux images  $I_1$  et  $I_2$  telles que :

$$G_i : \quad \begin{array}{l} I_i \longrightarrow I'_i \\ x_i \longrightarrow g_i(x_i) \end{array}$$

avec  $g_i(x_i)$  tel que :

$$\frac{\partial g_i(x_i)}{\partial Z} = K \quad (4)$$

pour  $i = 1, 2$ , où  $K$  est une constante choisie,  $x_i$  correspond à la coordonnée de l'image initiale. Définir la fonction  $g_i(x_i)$  telle que le rapport  $\partial g_i(x_i)/\partial Z$  soit constant correspond à imposer dans la nouvelle image que la

taille d'un pixel image (dans  $I'_1$  ou  $I'_2$ ) devienne invariante vis-à-vis d'une variation d'altitude. Sous cette contrainte, les pixels représentant des surfaces verticales (les façades) correspondront à des éléments de surface fixes et égaux dans les deux prises de vue. La transformation ainsi cherchée nous permettra de passer dans une image dont la géométrie sera comparable à celle présentée dans la figure 3 (b).

### 5.3 Etude de la transformation à appliquer aux images

Dans la suite, nous commencerons par nous intéresser à l'une des transformations que nous renommeons  $G : x \longrightarrow g(x)$ , appliquée à l'image  $I$  donnant l'image  $I'$ . L'équation de projection d'un point  $P(X, Z)$  de la scène bidimensionnelle  $S$  dans une ligne de l'image  $I$  est la suivante :

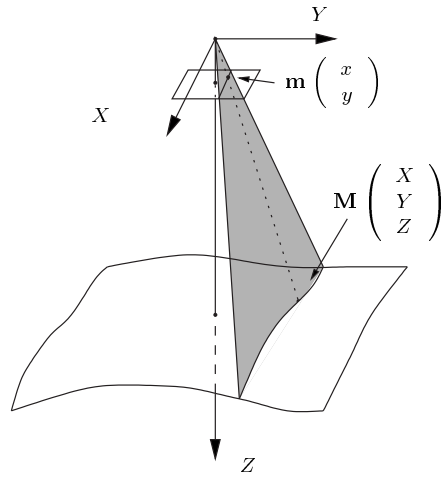


FIG. 5: *Modèle de projection sténopé en géométrie rectifiée (lignes épipolaires parallèles aux lignes images). Une droite épipolaire et les points de la scène 3D correspondant sont présentés.*

$$x = f_x \frac{X}{Z} + x_c$$

$f_x$  correspond au facteur multiplicatif selon l'axe des abscisses et  $x_c$  est l'abscisse du centre optique de l'image  $I$ . On recherche à présent l'application  $g(x)$  vérifiant la contrainte (4). Réécrivant cette contrainte sous forme décomposée et dérivant l'équation de projection, on obtient facilement l'équation suivante :

$$\begin{aligned} \frac{\partial g(x)}{\partial Z} &= \frac{\partial x}{\partial Z} \times \frac{\partial g(x)}{\partial x} \\ &= -f_x \frac{X}{Z^2} \times \frac{\partial g(x)}{\partial x} \\ &= K \end{aligned}$$

qui nous permet maintenant de redéfinir une nouvelle contrainte exprimée cette fois-ci vis-à-vis de l'abscisse  $x$  de l'image  $I$ .

$$\frac{\partial g(x)}{\partial x} = -\frac{K \times Z^2}{f_x \times X} \quad (5)$$

Reprenant l'équation de projection, on exprime l'abscisse  $X$  de l'espace bidimensionnel en fonction de l'altitude  $Z$  et de l'abscisse image  $x$  :

$$X = \frac{(x - x_c)Z}{f_x}$$

pour l'injecter dans la contrainte. On obtient enfin l'équation générale suivante permettant de définir la dérivée par rapport à  $x$  de la fonction de passage recherchée :

$$\frac{\partial g(x)}{\partial x} = -\frac{K Z(x)}{(x - x_c)} \quad (6)$$

Nous pouvons donc maintenant estimer cette fonction  $g(x)$  en intégrant l'équation 6. La fonction  $Z(x)$  correspond à la distance du sol au point  $x$ . Dans le cas général, il est nécessaire pour résoudre le problème d'estimer ou de linéariser cette fonction .

**Remarque** Dans le cadre de notre application, il convient de faire la distinction entre l'altitude et la variable  $Z$  que nous avons utilisée (cf figure 5).  $Z$  correspond à la distance qui sépare le sol du centre optique de notre caméra. L'altitude est, elle, déterminée à partir du sol.

**Hypothèse 2** *La projection de la scène dans le plan image est de type perspective faible*

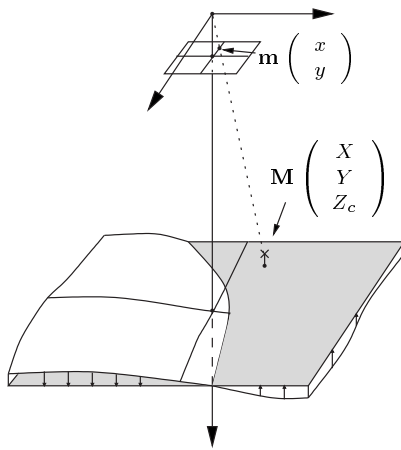


FIG. 6: *Modèle de projection sténopé en perspective faible. La scène est approximée par un plan.*

L'hypothèse 2 que nous venons d'énoncer nous permet de simplifier le problème par l'approximation de

la fonction de projection perspective. On suppose dans ce modèle que *la scène est un plan parallèle au plan image* comme le montre la figure 6. Dans cette condition, l'équation de projection devient simplement :

$$x = f_x \frac{X}{Z_c} + x_c$$

où  $Z_c$  est une constante. Tous les points de la scène possèdent donc la même distance  $Z_c$ . Cette projection se modélise en deux étapes successives : (1) projection des points de la scène sur le plan d'équation  $Z = Z_c$  puis (2) projection dans le plan image qui correspond dans ce cas précis à une homothétie de facteur  $f/Z_c$ . Reprenant la méthodologie précédemment employée on obtient la transformation  $g(x)$  suivante :

$$g(x) = -K Z_c \ln(x - x_c) \quad (7)$$

L'équation 7 correspondant au traitement à appliquer pour redresser les façades revient à rééchantillonner les images de manière logarithmique. Le résultat de cette "transformation logarithmique" est de "dilater" les zones de l'image en fonction de leur position par rapport au centre optique : les parties peu visibles (celles qui sont proches du centre optique) seront plus "dilatées" que les zones les plus visibles (celles qui sont les plus éloignées du centre optique : les bords des images). Les figures 7 et 8 nous présentent les résultats obtenus sur une paire d'images.

Le centre des images originales est peu utilisable dans un processus de corrélation de façades. Les façades sont très peu visibles et nous apportent donc une information moindre. Transformer cette partie de l'image pourrait se résumer à dilater quelques pixels pour représenter une façade. Il est donc inutile d'utiliser cette partie des images dans la configuration actuelle. Une solution serait d'utiliser des prises de vues obliques pour obtenir une meilleure visualisation des surfaces leur correspondant.

## 6 Etude du problème tridimensionnel

Dans cette section, nous traitons cas tridimensionnel. Nous allons nous intéresser aux surfaces. Dans un premier temps nous utilisons les résultats précédents pour généraliser la solution 2D trouvée au cas 3D. Nous allons ensuite proposer une nouvelle méthodologie pour traiter ce nouveau problème qui devient plus complexe. Nous nous intéresserons ensuite au cas du capteur sténopé avant de proposer une méthode adaptée à tout type de capteur.

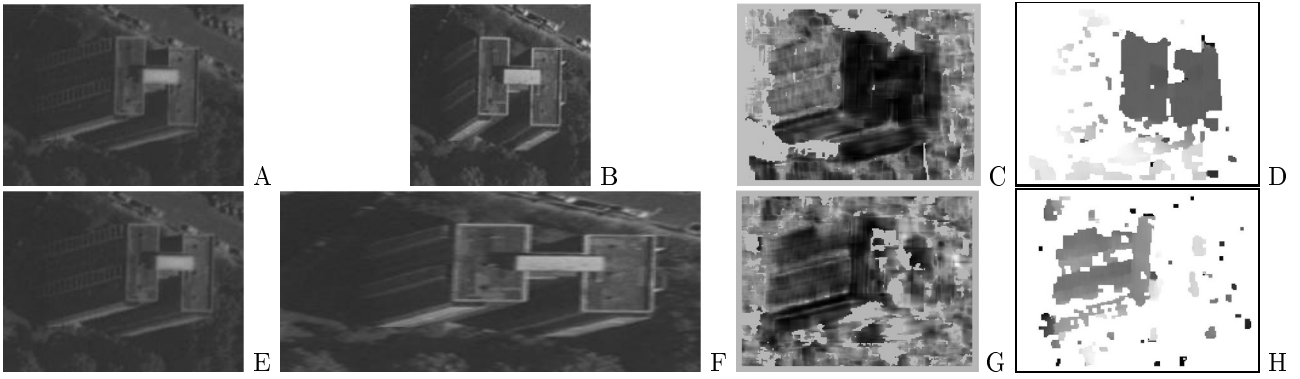


FIG. 7: Résultats obtenus avec la transformation calculée le long d'une courbe épipolaire. La première ligne (images A, B, C et D) présente les résultats d'appariement sans aucune transformation et la seconde (E, F, G et H) présente les résultats après avoir effectué notre transformation. A et B sont les images originales. C représente la carte des scores de corrélation pour ces deux images et D la carte de disparité correspondante après seuillage. On voit sur cette carte que seul le toit a "corrélé". E et F sont les images obtenues après transformation. On remarque que les façades sont de même taille. G et H sont respectivement la carte des scores de corrélation et la carte de disparité correspondante obtenues pour le couple E F. On voit que dans cette géométrie seule la façade a "corrélé".

## 6.1 Une première approche de la 3D

Avant de traiter le problème tridimensionnel dans sa généralité, nous utilisons les résultats précédents. Le passage de ce capteur à un capteur 3D peut être effectué alors de la façon suivante : on considère l'image comme étant la juxtaposition des droites épipolaires présentées auparavant mais en considérant cette fois qu'elles ne sont plus indépendantes.

La nouvelle transformation peut alors s'écrire sous la forme suivante :

$$G_i : \begin{array}{l} I_i \longmapsto I'_i \\ (x_i, y_i) \longmapsto (g_i(x_i), y_i) \end{array}$$

On souhaite obtenir une image dont la caractéristique essentielle est d'assurer un échantillonnage constant de la surface des façades. Le passage d'un point appartenant à la première image à la seconde dépend de sa position par rapport au centre optique. Un problème est posé alors pour les façades obliques (dont l'orientation n'est pas perpendiculaire à la direction des lignes de l'image). Les points appartenant à ces façades ne posséderont pas la même abscisse image, leurs projets vont donc être décalés le long des lignes comme le montre la figure 9.

Cet effet de déformation ou de rotation va dépendre de la position de la façade dans l'image. Il sera donc différent pour un même bâtiment en passant d'une image à l'autre. La mise en correspondance des façades sera donc perturbée puisque les voisinages des projetés d'un même point de cette façade ne se correspondront plus. Le cas 3D n'est donc pas déductible du cas 2D. Nous utiliserons donc dans la suite une nouvelle approche.

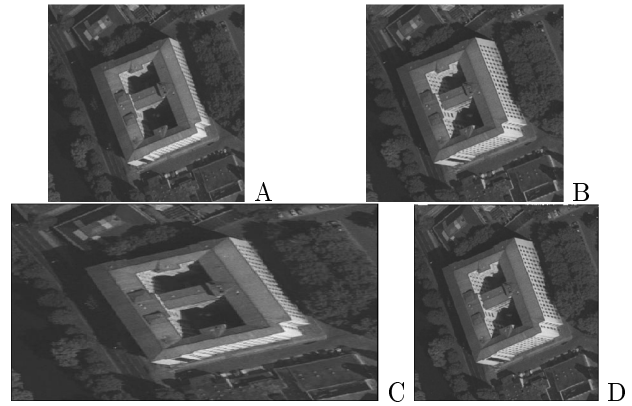


FIG. 9: Effet de rotation pour les façades dont l'orientation n'est pas perpendiculaire aux lignes images. A et B correspondent aux images initiales. Les images C et D correspondent à leurs transformées. On voit que les bâtiments sont sujets à un effet de rotation différent entre les deux images.





FIG. 8: Résultats obtenus en appliquant la transformation aux courbes épipolaires. A et B sont les deux images initiales et C et D les images reconstruites. Les façades sont perçues différemment dans les images initiales. Dans les images reconstruites, nous voyons que les façades des bâtiments dont l'orientation est perpendiculaire aux lignes sont à présent de surface équivalente. Les toits (correspondant à des surfaces horizontales) sont "étirés" et leurs surfaces sont perçues différemment. Nous pouvons voir aussi, dans le coin en haut à droite des images C et D, que les façades verticales, qui possèdent une orientation non orthogonale aux lignes des images, ne présentent pas les mêmes surfaces. La transformation n'a pas donné le résultat désiré pour leur cas.

## 6.2 Présentation du problème tridimensionnel

Nous souhaitons maintenant construire des images dans lesquelles chaque pixel représente un élément, 2D cette fois, régulier de façade. Un élément 2D caractérise maintenant un morceau de surface orientée (définie selon deux directions principales) dans la scène 3D. Nous traitons tout d'abord le cas général avec des images initiales résultantes d'un capteur générique. Nous cherchons cette fois à passer de cette classe d'images<sup>1</sup> à celle où nous pourrions mettre facilement les façades en correspondance. Nous nous intéresserons aussi à l'étude particulière du modèle sténopé orthographique.

## 6.3 Etude du cas général

Notre intérêt se porte à présent sur des éléments de surface orientés dans l'espace tridimensionnel. L'idée est de projeter des éléments de surface unitaires de la scène 3D dans les images par le modèle de projection générique. On recherche ensuite la transformation qui pourra "redresser" ces éléments, c'est-à-dire les remettre dans une géométrie adaptée à leur mise en correspondance. Comme nous l'avons vu, cette géométrie choisie est caractérisée par le fait que les surfaces de la scène 3D seront régulières dans ces images. C'est-à-dire qu'une façade d'orientation fixée sera perçue de façon similaire dans chacune des images.

On considère dans la suite notre scène comme une surface représentant le sol sur laquelle on place deux champs de vecteurs 3D constants : (1) le champ d'élevation représentant l'orientation de la verticale et (2) le champ d'orientation représentant l'orientation dans le plan horizontal. Ces deux champs de vecteurs orthogonaux sont générateurs de "façades" pour une orientation donnée. On mesure ensuite les champs projetés leur correspondant dans les images suivant le modèle de capteur (ou de projection) choisi. L'application consistant à "redresser" les façades ou surfaces verticales de la scène se résume alors à transformer les champs de vecteurs 2D obtenus par projection pour les rendre constants et égaux dans les deux images. Nous allons dans la suite présenter ces champs de vecteurs et déterminer leurs projetés. Nous nous intéresserons ensuite à la transformation permettant le passage entre les deux configurations de champs de vecteurs ainsi définis.

### 6.3.1 Etude des champs de vecteurs

Les éléments de surface 3D auxquels nous nous intéressons maintenant et dans la suite de ce document correspondent à des petits éléments 2D disposés sur la surface 3D représentant le sol notre scène. L'idée est de décomposer chacun d'entre eux sous la forme de

1. La classe d'images obtenue à partir d'un capteur générique.

deux éléments de variations directionnelles. On décide donc de décomposer chaque morceau de surface verticale orientée en deux vecteurs unitaires : (1) le vecteur d'élevation  $\vec{e}$  et (2) le vecteur d'orientation  $\vec{h}$ . On considère ainsi que pour chacun des éléments de surface verticale de notre scène correspond le même couple de vecteurs unitaires 3D  $(\vec{e}, \vec{h})$ . Notre scène est représentée alors comme la composée de deux champs de vecteurs 3D constants définis sur toute sa surface : (1) le champ d'élevation  $\vec{E}$  défini tel que pour tout point  $M(X, Y, Z)$  le vecteur  $\vec{e}_M$  du champ en ce point soit égal à :  $\vec{e} = (0, 0, 1)^T$ , et (2) le champ d'orientation  $\vec{H}$  défini tel que pour tout point  $M(X, Y, Z)$  le vecteur  $\vec{h}_M$  du champ en ce point soit égal à :  $\vec{h} = (X_h, Y_h, 0)^T$  avec  $X_h^2 + Y_h^2 = 1$ .

On s'intéresse maintenant à la projection de ces deux champs dans les images. On définit donc deux nouveaux champs de vecteurs 2D que nous nommerons  $\vec{p}_e$  et  $\vec{p}_h$  correspondant respectivement aux projetés des champs 3D  $\vec{E}$  et  $\vec{H}$  dans l'image par une projection quelconque  $Pr$ . On définit cette projection  $Pr$  d'un point  $M(X, Y, Z)$  de  $\mathcal{R}^3$  au point  $m(x, y)$  de  $\mathcal{R}^2$  de la manière suivante :

$Pr$  :

$$\begin{aligned} M &\longrightarrow m \\ (X, Y, Z) &\longrightarrow (x, y) = Pr(X, Y, Z) \end{aligned}$$

telle que

$$Pr(X, Y, Z) = (Px(X, Y, Z), Py(X, Y, Z))$$

Les champs 2D  $\vec{p}_e$  et  $\vec{p}_h$  correspondant aux projections des champs 3D  $\vec{E}$  et  $\vec{H}$  par  $Pr$  peuvent être déterminés de la manière suivante (cf figure 10) :

$$\vec{p}_e = \begin{pmatrix} x_{p_e}(x, y) \\ y_{p_e}(x, y) \end{pmatrix} = Jac(Pr)\vec{E} \quad (8)$$

$$\vec{p}_h = \begin{pmatrix} x_{p_h}(x, y) \\ y_{p_h}(x, y) \end{pmatrix} = Jac(Pr)\vec{H} \quad (9)$$

où  $Jac(Pr)$  est le jacobien de la projection  $Pr$  et est défini comme suit :

$$Jac(Pr) = \begin{pmatrix} \frac{\partial Px}{\partial X}(X, Y, Z) & \frac{\partial Px}{\partial Y}(X, Y, Z) & \frac{\partial Px}{\partial Z}(X, Y, Z) \\ \frac{\partial Py}{\partial X}(X, Y, Z) & \frac{\partial Py}{\partial Y}(X, Y, Z) & \frac{\partial Py}{\partial Z}(X, Y, Z) \end{pmatrix}$$

Les deux champs  $\vec{p}_e$  et  $\vec{p}_h$  sont des champs bidimensionnels, définis dans l'image  $I$ , dont les composantes vont dépendre du modèle de projection  $Pr$  utilisé puisque les champs 3D  $\vec{E}$  et  $\vec{H}$  sont constants et connus. On souhaite maintenant les transformer pour obtenir une nouvelle image  $I'$  dans laquelle on va contraindre leurs images respectives. Calculons la transformation affine

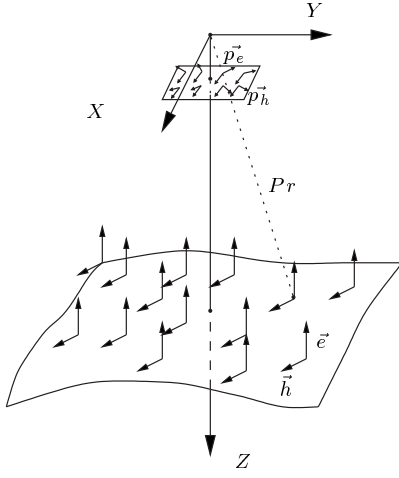


FIG. 10: Présentation des champs 3D de la scène dans l'image.  $\vec{E}$  est le champ d'élévation, il est représenté en tout point par le vecteur  $\vec{e}$ . Le champ 3D d'orientation  $\vec{H}$  est représenté en tout point par le vecteur  $\vec{h}$ . Leurs projections dans l'image correspondent aux champs de vecteurs 2D  $\vec{p}_e$  et  $\vec{p}_h$ .

$T_{pq}$  du repère  $(\vec{p}_e, \vec{p}_h)$  au point  $(x, y)$  de l'image  $I$  au repère  $(\vec{q}_e, \vec{q}_h)$  au point  $(u, v)$  lui correspondant dans la nouvelle image  $I'$  comme le présente la figure 11 (les points  $(x, y)$  de  $I$  et  $(u, v)$  de  $I'$  sont les projections du même point 3D de la scène dans les deux images) :

$$\begin{aligned} T_{pq} &: I \longrightarrow I' \\ \vec{p}_e &\longrightarrow \vec{q}_e = T_{pq}(\vec{p}_e) \\ \vec{p}_h &\longrightarrow \vec{q}_h = T_{pq}(\vec{p}_h) \end{aligned}$$

que l'on peut réécrire sous la forme suivante

$$\begin{aligned} \begin{pmatrix} x_{q_e}(u, v) \\ y_{q_e}(u, v) \end{pmatrix} &= \begin{pmatrix} Ta(x, y) & Tb(x, y) \\ Tc(x, y) & Td(x, y) \end{pmatrix} \begin{pmatrix} x_{p_e}(x, y) \\ y_{p_e}(x, y) \end{pmatrix} \\ \begin{pmatrix} x_{q_h}(u, v) \\ y_{q_h}(u, v) \end{pmatrix} &= \begin{pmatrix} Ta(x, y) & Tb(x, y) \\ Tc(x, y) & Td(x, y) \end{pmatrix} \begin{pmatrix} x_{p_h}(x, y) \\ y_{p_h}(x, y) \end{pmatrix} \end{aligned}$$

La transformation  $T$  cherchée a quatre inconnues. Si l'on définit les quatre champs  $\vec{p}_e$  et  $\vec{p}_h$  (évalués tous deux à partir de la projection  $Pr$ ) ainsi que  $\vec{q}_e$  et  $\vec{q}_h$  que l'on choisit, on dispose alors d'un système de quatre équations nous permettant de déterminer toutes les composantes de la transformation  $T_{pq}$ . Connaissant les champs 2D  $\vec{p}_e$ ,  $\vec{p}_h$ ,  $\vec{q}_e$  et  $\vec{q}_h$  en tous point  $(x, y)$  de l'image  $I$ , on peut calculer alors les valeurs de toutes les composantes de la transformation étudiée  $T_{pq}$  :

$$\begin{aligned} Ta &= \frac{y_{p_h} x_{q_e} - y_{p_e} x_{q_h}}{x_{p_e} y_{p_h} - y_{p_e} x_{p_h}} & Tb &= \frac{x_{p_e} x_{q_h} - x_{p_h} x_{q_e}}{x_{p_e} y_{p_h} - y_{p_e} x_{p_h}} \\ Tc &= \frac{y_{p_h} y_{q_e} - y_{p_e} y_{q_h}}{x_{p_e} y_{p_h} - y_{p_e} x_{p_h}} & Td &= \frac{x_{p_e} y_{q_h} - x_{p_h} y_{q_e}}{x_{p_e} y_{p_h} - y_{p_e} x_{p_h}} \end{aligned}$$

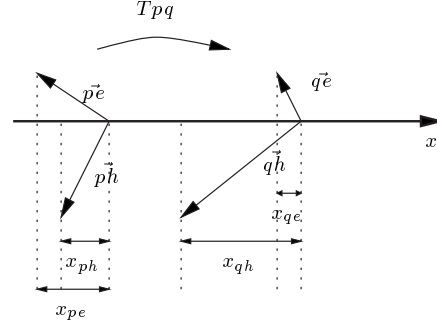


FIG. 11: Présentation de la transformation  $T_{pq}$  permettant de passer du repère correspondant aux champs de vecteurs 2D  $(\vec{p}_e, \vec{p}_h)$  au repère correspondant aux champs de vecteurs 2D  $(\vec{q}_e, \vec{q}_h)$ .

où chaque composante des champs 2D  $\vec{p}_e$  et  $\vec{p}_h$  et chaque composante de la transformation  $T_{pq}$  sont des fonctions dépendants de  $x$  et de  $y$ .

Nous avons vu précédemment que le couple de vecteurs  $(\vec{p}_e, \vec{p}_h)$  défini en un point  $(x, y)$  représente la projection dans l'image  $I$  d'un élément de surface vertical de la scène 3D. On utilise maintenant la transformation  $T_{pq}$  ainsi définie pour associer chaque élément de surface 2D de l'image de départ  $I$  à celui qui lui correspond dans la nouvelle image  $I'$ . Cette transformation est donc une application qui associe à chaque point  $m(x, y)$  de l'image  $I$  le même point  $m(u, v)$  de l'image  $I'$  de la manière suivante :

$$\begin{pmatrix} du \\ dv \end{pmatrix} = \begin{pmatrix} Ta & Tb \\ Tc & Td \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} \quad (10)$$

où  $(du, dv)$  représente un élément de surface de l'image  $I'$  et  $(dx, dy)$  un élément de surface de l'image  $I$ . L'application  $T_{pq}$  ainsi définie nous permet donc (si l'on connaît les champs 2D projetés  $\vec{p}_e$ ,  $\vec{p}_h$ ,  $\vec{q}_e$  et  $\vec{q}_h$ ) d'associer les deux images  $I$  et  $I'$ . On peut la décomposer en deux systèmes de deux équations à deux inconnues en séparant les deux variables  $u$  et  $v$ . L'application  $T_{pq}$  est donc définie par le système différentiel suivant :

$$du = Ta dx + Tb dy \quad (11)$$

$$dv = Tc dx + Td dy \quad (12)$$

L'idée est maintenant de préserver la géométrie épipolaire du système. On cherche donc une transformation  $T_{pq}$  qui s'opère le long des lignes images. On impose ainsi la contrainte  $dv = dy$ . Celle-ci se répercute uniquement dans la seconde équation et impose :  $Tc = 0$  et  $Td = 1$ . Le problème posé maintenant est de résoudre la première équation (11). Cette équation est une forme différentielle. La résoudre équivaut à cher-

cher une fonction  $U(x, y)$  telle que :

$$\begin{aligned}\frac{\partial U(x, y)}{\partial x} &= Ta(x, y) \\ \frac{\partial U(x, y)}{\partial y} &= Tb(x, y)\end{aligned}$$

avec

$$\frac{\partial Ta(x, y)}{\partial y} - \frac{\partial Tb(x, y)}{\partial x} = 0 \quad (13)$$

Sous ces conditions, nous pourrions affirmer que la fonction  $U(x, y)$  est une intégrale première de la forme différentielle (11). On ne traitera pas ici les cas où la solution ne vérifie pas l'égalité de Schwarz (13). La solution finale sera donc de la forme :

$$\begin{cases} u = U(x, y) \\ v = y \end{cases} \quad (14)$$

## 6.4 Etude avec un capteur sténopé orthographique

On étudie à présent le cas du modèle de projection sténopé orthographique<sup>2</sup>. Nous cherchons à obtenir une solution direction au problème considéré. Nous commençons par calculer les valeurs des champs d'élévations et d'orientation 2D projetés par ce modèle de projection. La scène 3D étant planaire, nous obtenons les champs suivant :

$$\begin{cases} \vec{p}_e = \begin{pmatrix} x_{p_e}(x, y) \\ y_{p_e}(x, y) \end{pmatrix} = \begin{pmatrix} \frac{x}{Z_c} \\ \frac{y}{Z_c} \end{pmatrix} \\ \vec{p}_h = \begin{pmatrix} x_{p_h}(x, y) \\ y_{p_h}(x, y) \end{pmatrix} = \begin{pmatrix} \frac{X_h}{Z_c} \\ \frac{Y_h}{Z_c} \end{pmatrix} = \begin{pmatrix} x_{h1} \\ y_{h1} \end{pmatrix} \end{cases}$$

Ces deux champs sont présentés dans la figure 12. Les composantes du premier champ,  $p_{\vec{E}}$  sont des fonctions de  $x$  et  $y$  tandis que les composantes du second sont constantes. On utilise alors les fonctions définies ci-dessus pour calculer les coefficients de la forme différentielle du système que nous voulons résoudre :

$$Ta(x, y) = \frac{Z_c y_{h1} x_{q_e} - y \times x_{q_h}}{x \times y_{h1} - y \times x_{h1}} \quad (15)$$

$$Tb(x, y) = \frac{x \times x_{q_h} - Z_c y_{h1} x_{q_e}}{x \times y_{h1} - y \times x_{h1}} \quad (16)$$

L'égalité de Schwarz n'est pas vérifiée pour ce modèle de projection dans le cas général où l'on décide d'imposer  $x_{q_h} \neq 0$  et  $x_{q_e} \neq 0$ . On choisit de traiter tout d'abord le cas où  $x_{q_h} = 0$  et  $x_{q_e} \neq 0$ . Ce cas de figure correspond à se transporter dans une image où seule l'élévation varie le long des lignes épipolaires. Les déplacements

<sup>2</sup>. Ceci nous permet de faire le lien avec les résultats obtenus pour le problème bidimensionnel (cf section 5).

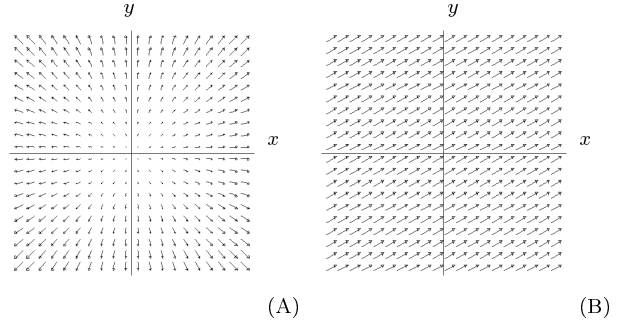


FIG. 12: Champs de vecteurs représentant les champs de vecteurs 2D d'élévation  $\vec{p}_e$  (projection du champ 3d  $\vec{E}$ )(A) et d'orientation  $q_e$  (projection du champ 3d  $\vec{H}$ )(B) obtenus avec un modèle de capteur sténopé orthographique.

sur les façades 3D le long du champ d'orientation n'entraîneront ainsi aucun déplacement le long des lignes. C'est donc un cas particulier dans lequel toutes les surfaces étudiées sont réorientées perpendiculairement aux lignes de l'image.

La forme différentielle étant exacte, nous l'intégrons directement pour obtenir la solution suivante :

$$u(x, y) = x_{q_e} Z_c \ln(-x \times y_{h1} + y \times x_{h1}) \quad (17)$$

On remarque que la solution obtenue ici (17) reprend la forme de la solution bidimensionnelle (équation 7) lorsque l'on choisit une orientation initiale des façades rendant le vecteur d'orientation  $\vec{p}_h = (x_{h1}, y_{h1})$  perpendiculaire aux lignes épipolaires (donc  $x_{h1} = 0$  et  $y_{h1} \neq 0$ ). La figure 13 présente les résultats obtenus en utilisant cette solution.

Le problème n'a pas de solution directe lorsque nous choisissons de construire des images avec un champ final d'orientation  $\vec{p}_h$  tel que  $x_{p_h}$  soit non nul. Nous proposons de résoudre le problème en le linéarisant.

## 6.5 Linéarisation du problème

Dans ce qui précède, nous supposons connaître le modèle de projection de manière littérale. Dans la réalité, c'est rarement le cas et aucune solution directe n'est disponible. On résout le problème localement en linéarisant la transformation. Il nous faut estimer les deux valeurs des coefficients  $Ta$  et  $Tb$  au point  $(x_0, y_0)$  étudié et dans son voisinage. La linéarisation de la projection  $Pr$  nous ramène à estimer les valeurs en tout point de l'image du jacobien de cette transformation. Reprenons la projection  $Pr$  définie comme suit :

$$(x, y) = Pr(X, Y, Z) = (Px(X, Y, Z), Py(X, Y, Z))$$

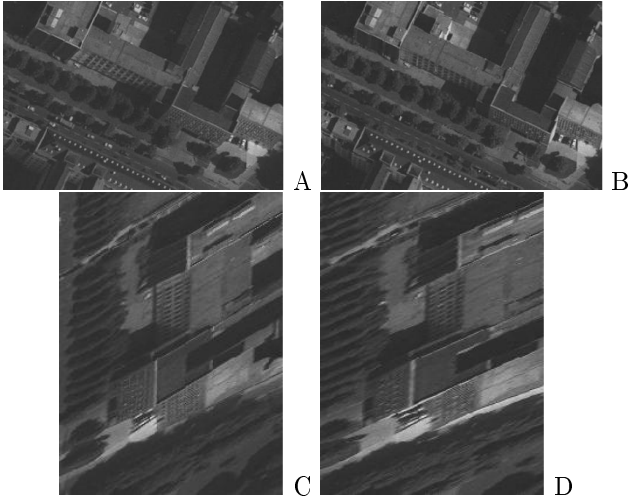


FIG. 13: Résultats obtenus dans le cas où nous choisissons  $x_{q_h} = 0$  et  $x_{q_e} \neq 0$ . A et B sont les images originales et C et D sont les images obtenues après transformation. On peut voir que les façades sont "ré-orientées" perpendiculairement aux droites épipolaires.

En choisissant un voisinage suffisamment petit, on linéarise la projection  $Pr$  au point  $(X_0, Y_0, Z_0)$  de la façon suivante :

$$(x, y) = P(X_0, Y_0, Z_0) + (X - X_0) \times \left. \frac{\partial P}{\partial X} \right|_{X_0, Y_0, Z_0} + (Y - Y_0) \times \left. \frac{\partial P}{\partial Y} \right|_{X_0, Y_0, Z_0} + (Z - Z_0) \times \left. \frac{\partial P}{\partial Z} \right|_{X_0, Y_0, Z_0}$$

qui est équivalent au système suivant :

$$(x, y) = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{pmatrix} + (x_0, y_0)$$

tel que :

$$(x_0, y_0) = Pr(X_0, Y_0, Z_0)$$

$$\begin{aligned} a &= \left. \frac{\partial Px}{\partial X} \right|_{X_0, Y_0, Z_0} & b &= \left. \frac{\partial Px}{\partial Y} \right|_{X_0, Y_0, Z_0} \\ c &= \left. \frac{\partial Px}{\partial Z} \right|_{X_0, Y_0, Z_0} & d &= \left. \frac{\partial Py}{\partial X} \right|_{X_0, Y_0, Z_0} \\ e &= \left. \frac{\partial Py}{\partial Y} \right|_{X_0, Y_0, Z_0} & f &= \left. \frac{\partial Py}{\partial Z} \right|_{X_0, Y_0, Z_0} \end{aligned}$$

Nous allons donc utiliser ces équations pour estimer le jacobien  $Jac(Pr)$  de la transformation au point  $(x_0, y_0)$ . Dans notre cas, nous nous intéressons à la projection des champs de vecteurs 3D d'élévation et

d'orientation représentant nos surfaces. On peut estimer maintenant les valeurs des champs 2D correspondant à leur projection à l'aide des équations suivantes :

$$\vec{p}_e(x_0, y_0) = \begin{pmatrix} x_{p_e}(x_0, y_0) \\ y_{p_e}(x_0, y_0) \end{pmatrix} = \begin{pmatrix} c \\ f \end{pmatrix} \quad (18)$$

$$\begin{aligned} \vec{p}_h(x_0, y_0) &= \begin{pmatrix} x_{p_h}(x_0, y_0) \\ y_{p_h}(x_0, y_0) \end{pmatrix} \\ &= \begin{pmatrix} a & b \\ e & f \end{pmatrix} \begin{pmatrix} X_h \\ Y_h \end{pmatrix} \end{aligned} \quad (19)$$

qui correspondent à deux champs 2D constants dans le voisinage considéré autour du point  $(x_0, y_0)$ . Nous pouvons maintenant calculer les valeurs des coefficients  $Ta$  et  $Tb$  au point  $(x_0, y_0)$  de la forme différentielle et obtenir la transformation  $T_{pq}(x_0, y_0)$  locale au même point :

$$u(x, y) = Ta(x_0, y_0) x + Tb(x_0, y_0) y \quad (20)$$

avec

$$\begin{aligned} Ta(x_0, y_0) &= \frac{(dX_h + eY_h)x_{q_e} - fx_{q_h}}{(cd - fa)X_h + (ce - fb)Y_h} \\ Tb(x_0, y_0) &= \frac{cx_{q_h} - (aX_h + bY_h)x_{q_e}}{(cd - fa)X_h + (ce - fb)Y_h} \end{aligned}$$

La figure 14 présente les résultats obtenus. Le choix des paramètres joue un rôle important dans la qualité des résultats.  $X_h$  et  $Y_h$  définissent l'orientation des façades étudiées. Nous les choisissons conformément à la façade étudiée (dans le cadre d'une approche locale). Le choix des composantes  $x_{q_e}$  et  $x_{q_h}$  s'effectue de façon à limiter les déformations du rééchantillonnage. Dans le cas de deux vues, nous pouvons choisir de reconstruire nos images en imposant aux valeurs des champs 2D  $\vec{q}_e$   $\vec{q}_h$  celles des champs  $\vec{p}_e$  et  $\vec{p}_h$  estimés pour le modèle cyclopéen.

## 7 Résultats de reconstruction

Nous nous intéressons dans cette section à la mise en correspondance et à la reconstruction des façades de bâtiments. Une fois les images "redressées" construites, nous appliquons un algorithme de corrélation standard avec le critère du ZNCC (cf [6]) sur une fenêtre de taille (11, 7). Nous filtrons ensuite la carte de disparité puis reconstruisons les points 3D correspondant à chaque appariement en revenant dans les images initiales par la transformation inverse.

Deux résultats sont présentés dans les figures 15 et 16. Les résultats présentés dans la première figure permettent de vérifier tout d'abord que l'orientation est bien prise en compte et qu'il n'y a plus d'effet de rotation sur le bâtiment. Les points reconstruits s'accumulent le long des façades du bâtiment. Les résultats présentés dans la seconde figure sont similaires.



FIG. 14: Résultats obtenus en estimant la transformation locale. *A* et *B* sont les deux images initiales. *C* et *D* sont les images reconstruites avec  $x_{q_e} = 1$  et  $x_{q_h} = 2$  (Ce choix des coefficients permet de limiter les déformations introduites par le rééchantillonnage). Les façades étudiées sont celles dont l'orientation est de  $115^\circ$  par rapport aux lignes. Dans les images initiales, ces façades sont perçues différemment et ne peuvent être appariées. Dans les images reconstruites, nous voyons qu'elles sont à présent dans une même géométrie.

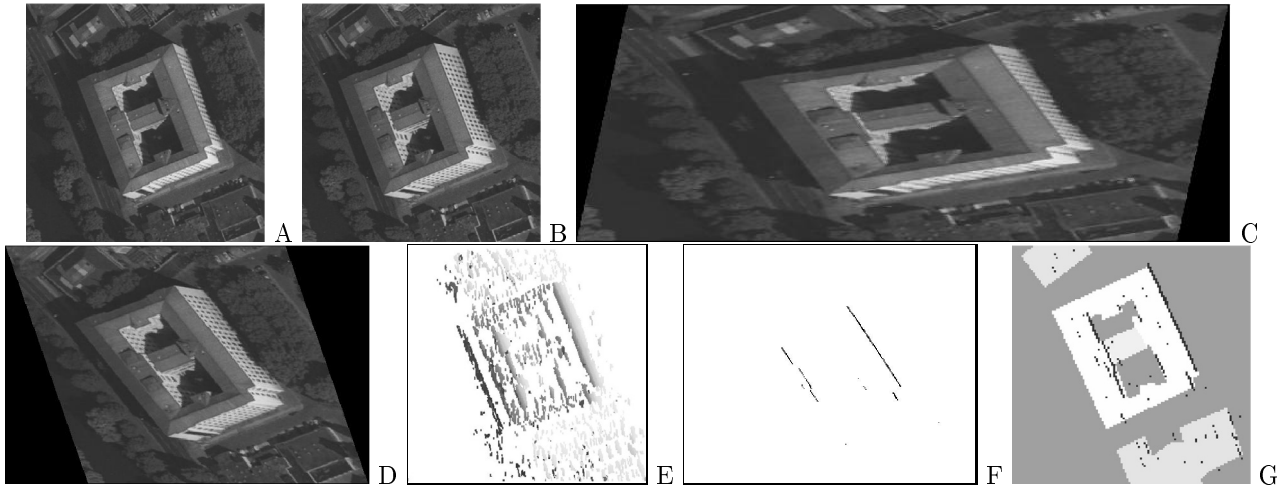


FIG. 15: Résultats de reconstruction d'un bâtiment. A et B sont les images originales. C et D sont respectivement les transformées des images A et B avec  $x_{q_h} = x_{q_e} = 2$ . E est la carte de disparité reprojétée en géométrie originale (projection de la carte de disparité obtenue après mise en correspondance des images C et D). F est un résultat de reconstruction 2,5D filtré (résolution de 50cm). Chaque point 2,5D de l'image représente un point d'accumulation des appariements trouvés. G est le résultat projeté dans le MNE correspondant à la zone étudiée (résolution de 1m). On voit que ces droites correspondent bien aux bords des façades.

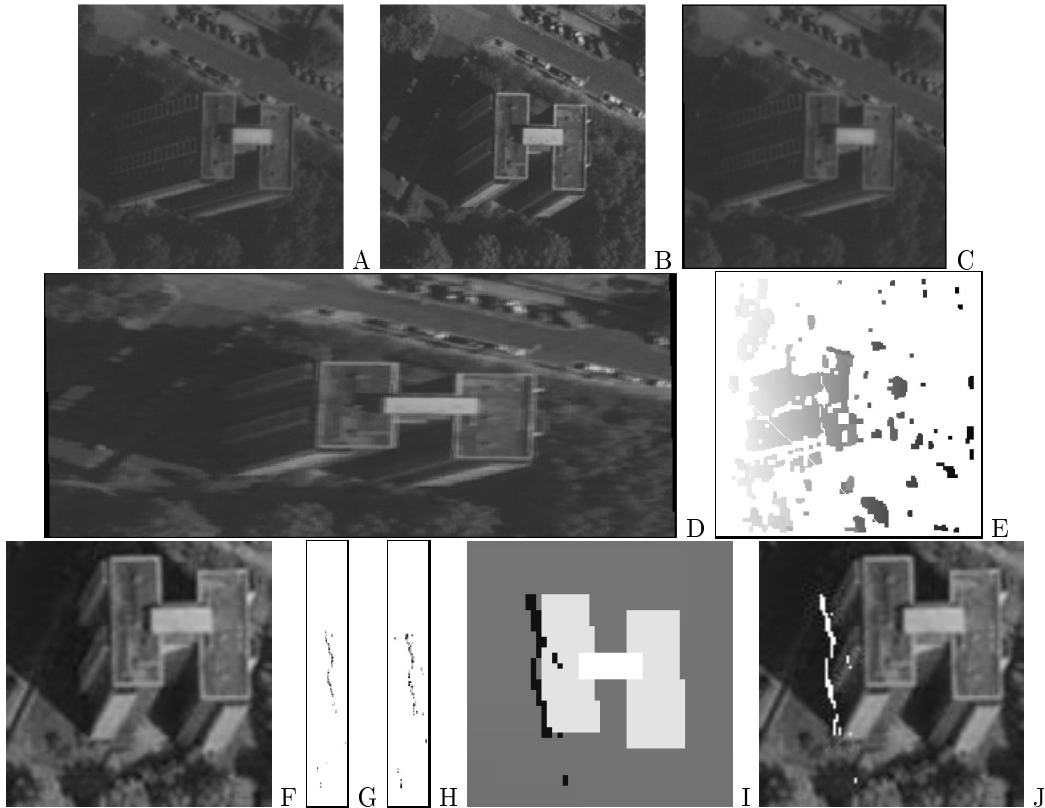


FIG. 16: Résultats de reconstruction d'un bâtiment. A et B sont les images originales. C et D sont respectivement les transformées des images A et B avec  $x_{q_h} = x_{q_e} = 2$ . E est la carte de disparité reprojétée en géométrie originale. F est une ortho-image. G et H sont deux résultats de reconstruction 2,5D filtrés (résolution de 25cm et 50cm). I est le résultat projeté dans le MNE correspondant à la zone étudiée (résolution de 1m). On retrouve les formes des balcons. J est la projection du résultat dans l'ortho-image.

La forme de la façade (les balcons) est retrouvée. Les essais nous ont montré que le choix de la résolution de reconstruction est très importante. La précision obtenue est inférieure au pixel puisque les points apparés dans les nouvelles images sont rétroprojetés avec une précision "sous-pixelique". Une étude qualitative de cette précision est envisagée.

## 8 Conclusion

Nous avons présenté dans cet article un nouveau concept permettant d'utiliser beaucoup plus facilement les informations apportées par les images des façades des bâtiments de la scène 3D. Pour cela, nous appliquons le "redressement de façades" à chacune des images. Les nouvelles images obtenues possèdent une géométrie simplifiant la mise en correspondance des points 3D appartenant aux façades. Elles peuvent dès lors être utilisées dans un processus d'appariement et de reconstruction automatique ou semi-automatique avec l'aide d'un opérateur externe.

En utilisant ce concept, deux approches peuvent être choisies pour reconstruire les façades. L'approche "globale" est d'appliquer le redressement de chaque image pour toutes les orientations possibles des façades. On recherche ensuite les points se correspondant pour retrouver toutes les surfaces verticales de la scène. Cette approche est sans aucun doute impraticable du fait de l'immensité de l'espace de recherche et des fortes distorsions introduites par un rééchantillonnage abusif. L'approche dite "locale" consiste, elle, à estimer le redressement local propre à chaque façade des bâtiments de la scène pour retrouver la position exacte de celles-ci. Il est nécessaire alors de bien choisir les champs de vecteurs finaux imposés pour limiter les effets de distorsion dus aux rééchantillonnages de chacune des images. Cette seconde approche est celle que nous avons adoptée pour les développements futurs.

Les résultats de reconstruction présentés sont très prometteurs. Nous nous intéressons maintenant à l'étude qualitative des façades reconstruites ainsi qu'à la réalisation d'un processus de reconstruction automatique de bâtiment incluant le "redressement de façades".

## Références

- [1] J.Y. Aloimonos. Perspective approximations. *Image and Vision Computing*, 8(3):179–192, August 1990.
- [2] Caroline Baillard. *Analyse d'images aériennes stéréoscopiques pour la restitution 3-D des milieux urbains*. PhD thesis, École Nationale Supérieure des Télécommunications, October 97.
- [3] Frank Bignone, Olof Henricsson, Pascal Fua, and Markus Stricker. Automatic extraction of generic house roofs from high resolution aerial imagery. In Bernard Buxton, editor, *Proceedings of the 4th European Conference on Computer Vision*, pages 85–96, Cambridge, UK, April 1996.
- [4] Robert Collins, Chris Jaynes, Yong-Qing Cheng, Xiaoguang Wang, Frank Stolle, Howard Schultz, Allen Hanson, and Edward Riseman. The umass ascender system for 3d site model construction. In *RADIUS: Image Understanding for Imagery Intelligence*, pages 209–222, 1996.
- [5] Matthieu Cord. *Analyse d'images aériennes haute résolution : détection et modélisation du bâti en zone urbaine*. PhD thesis, Université de Cergy-Pontoise, ENSEA-ETIS, December 1998.
- [6] Frédéric Devernay. *Vision stéréoscopique et propriétés différentielles des surfaces*. PhD thesis, École Polytechnique, February 97.
- [7] Umesh R. Dhond and J.K. Aggarwal. Structure from stereo - a review. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(6):1489–1510, 1989.
- [8] O. Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, 1993.
- [9] Laurent Gabet, Gérard Giraudon, and Laurent Renouard. Construction automatique de modèle numérique de terrain haute résolution en zone urbaine. In *Bulletin S.F.T.P.*, volume 135, pages 9–25, 1994.
- [10] Stéphane Girard, Philippe Guérin, Henri Maître, and Michel Roux. Construction d'une carte de disparité dense pour la reconnaissance de bâtiments en imagerie aérienne. In *11ème Congrès RFIA '98*, pages 319–328, 1998.
- [11] R. Hartley and R. Gupta. Linear pushbroom cameras. In J-O. Eklundh, editor, *Proceedings of the 3rd European Conference on Computer Vision*, volume 800-801 of *Lecture Notes in Computer Science*, Stockholm, Sweden, May 1994. Springer-Verlag.
- [12] Olof Henricsson. *Analysis of image structures using color attributes and similarity relations*. PhD thesis, Swiss Federal Institute of Technology Zurich, November 96.
- [13] Stéphane Laveau. *Géométrie d'un système de N caméras. Théorie, estimation et applications*. PhD thesis, École Polytechnique, May 96.
- [14] Jean-Luc Lotti. *Mise en correspondance stéréo par fenêtres adaptatives en imagerie haute résolution*. PhD thesis, Université de Nice - Sophia Antipolis, France, February 1996.
- [15] Henri Maître and Wei Luo. Using models to improve stereo reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):269–277, February 1992.
- [16] Theo Moons, David Frère, Jan Vandekerckhove, and Luc Van Gool. Automatic modelling and 3d reconstruction of urban house roofs from high resolution aerial imagery. In *Proceedings of the 5th European Conference on Computer Vision*, pages 410–425, Freiburg, Germany, June 98.
- [17] Michel Roux, Yuan Hsieh, and David McKeown. Performance analysis of object space matching for building extraction using several images. In SPIE, editor, *Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision II*, volume 2486, pages 277–297, Orlando, Florida, 1995.
- [18] Uwe Weidner and Wolfgang Förstner. Toward automatic building extraction from high resolution digital elevation models. In *ISPRS Journal of Photogrammetry and Remote Sensing*, volume 50, pages 38–49, 1995.