Introduction

Notre objectif est de pour réaliser un panorama à l'aide de deux images différentes. Ces images ont en réalité des points qui correspondent à l'autre image qui peuvent avoir été obtenu manuellement ou par un autre algorithme. Le principe de la méthode est de transformer une des images dans le repère de l'autre à l'aide des points correspondants. Nous nous focaliserons d'abord sur une partie théorie puis sur la transformation avec des points choisi manuellement, ensuite nous verrons l'algorithme RANSAC qui permet de manipuler les données. Cet algorithme est nécessaire car dans le cas de point donné par un algorithme il peut y avoir des points faux qui fausseront complètement la transformation.

Théorie transformation dans le repère de l'image 2

Une homographie peut être interprété comme une composition de translation, rotation et homothétie. Comme expliqué dans l'Appendix trouver H tel que :

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = H \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \text{ (1)}$$

En effet H peut être interpréter comme un changement de repère entre le plan image de l'image que l'on cherche à transformer dans le repère camera de l'image dans lequel on veut projeter.

Ainsi $\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$ représente les coordonnées homogènes connu du pixel de l'image 1 dans son propre repère « image » et $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$ les coordonnée 3D du point désigné par le pixel de l'image 1 dans le repère « caméra » de l'image 2. Nous connaissons aussi $\begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}$ les coordonnées homogènes du pixel de l'image 2 dans le repère « Image » de l'image 2, celui-ci est associé au point $\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$ de l'image 1.

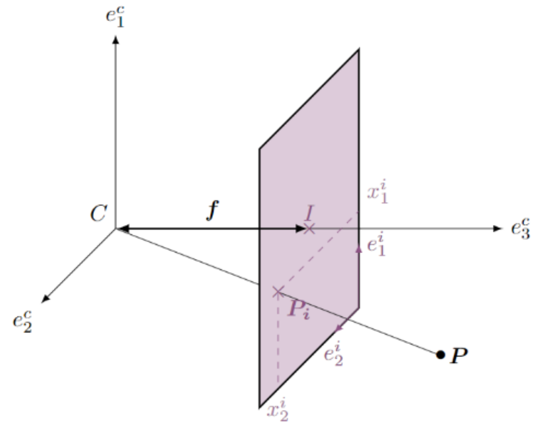


FIGURE 4 – Passage du repère caméra au repère image

[1]

Figure 1 : repère camera et image de l'image 2 (plan dans lequel on veut projeter l'image 1)

On peut visualiser sur le schéma ci-dessus que $P\begin{pmatrix}x'\\y'\\z'\end{pmatrix}$ sont les coordonnées du point résultant de la transformé H du pixel de l'image 1. On observe aussi que pour ce point 3D, nous connaissons sa projection $Pi\begin{pmatrix}X\\Y\\1\end{pmatrix}$ car c'est le point qui correspond au pixel choisi dans l'image 1

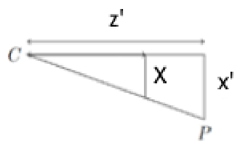


FIGURE 5 – Relation entre coordonnées des repères caméra et image

On obtient donc d'après le théorème de Thales $X = \frac{x'}{z'}$ de même on trouve $Y = \frac{y'}{z'}$

Par identification avec l'équation 1 nous obtenons :

$$X = \frac{x'}{z'} = \frac{H_{11}x + H_{12}y + H_{13}}{H_{31}x + H_{32}y + H_{33}} \quad (2)$$

$$Y = \frac{y'}{z'} = \frac{H_{21}x + H_{22}y + H_{33}}{H_{31}x + H_{32}y + H_{33}} \quad (3)$$

[1] Note : le repère image dans notre cas n'est pas au centre de l'image comme représenté dans le schéma

Avec $H = \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{32} & H_{32} & H_{33} \end{pmatrix}$

On réécrit ces équations sous la forme matriciel avec :

$$h = (H_{11}, H_{12}, H_{13}, H_{21}, H_{22}, H_{23}, H_{31}, H_{32}, H_{33})$$

Et pour chaque couple de point que l'on dispose, nous définissons donc les vecteurs :

$$r_x = (x \quad y \quad 1 \quad 0 \quad 0 \quad 0 \quad -Xx \quad -Xy \quad -X)$$

$$r_y = (0 \quad 0 \quad 0 \quad x \quad y \quad 1 \quad -Yx \quad -Yy \quad -Y)$$

En posant n le nombre de couple de point correspondant entre les deux images, nous obtenons un système d'équation à 2n équations que nous mettons sous la forme matricielle :

$$M.h^T = \begin{pmatrix} r_x^1 \\ r_y^1 \\ \vdots \\ r_x^n \\ r_y^n \end{pmatrix} h^T = 0$$

Or cette équation ne possède pas de solutions exactes, nous trouvons donc la solution qui est la plus proches au sens des moindres carré. En effet la résolution au sens des moindres carrée peut être réalisé par une décomposition en SVD et la solution h sera un vecteur propre associé à la plus petite des valeurs singulières. En effet la dernière valeur singulière est la valeur singulière qui possède le moins d'énergie et donc minimise le produit $M.h^T$

Une fois que nous retrouvons la transformé H il suffit de projeter les points de l'image 1 dans le repère image 2 à l'aide des formules (2) et (3)

$$X = \frac{H_{11}x + H_{12}y + H_{13}}{H_{31}x + H_{32}y + H_{33}} \;\; et \;\; Y = \frac{H_{21}x + H_{22}y + H_{33}}{H_{31}x + H_{32}y + H_{33}}$$

Et obtenir la transformé de l'image 1 dans le repère image de l'image 2.

Première résolution

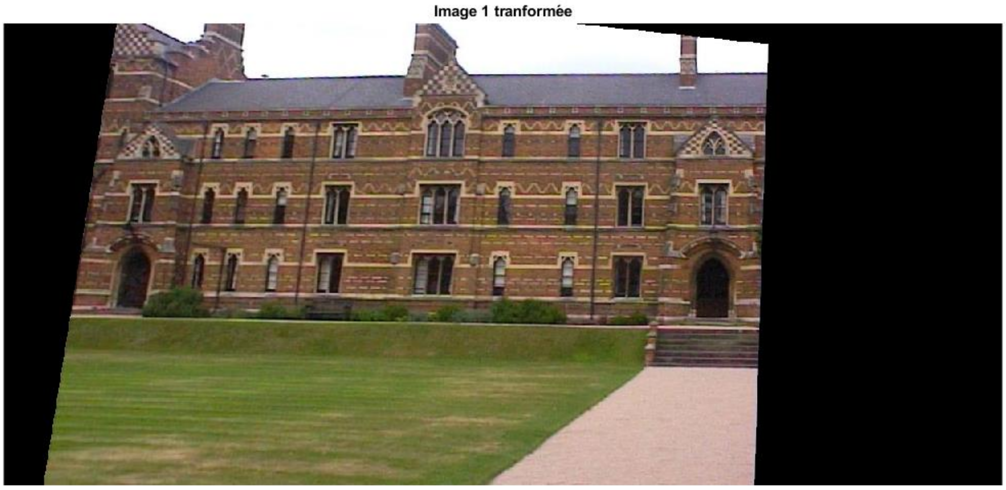
Quatre points correspondants entre l'image 1 et 2 ont été choisis manuellement pour être sûr de leurs fiabilités. Nous cherchons donc connaissant ces 4 points l'homographie H qui permettra de mettre l'image dans la même disposition que l'image 2.

Image 1

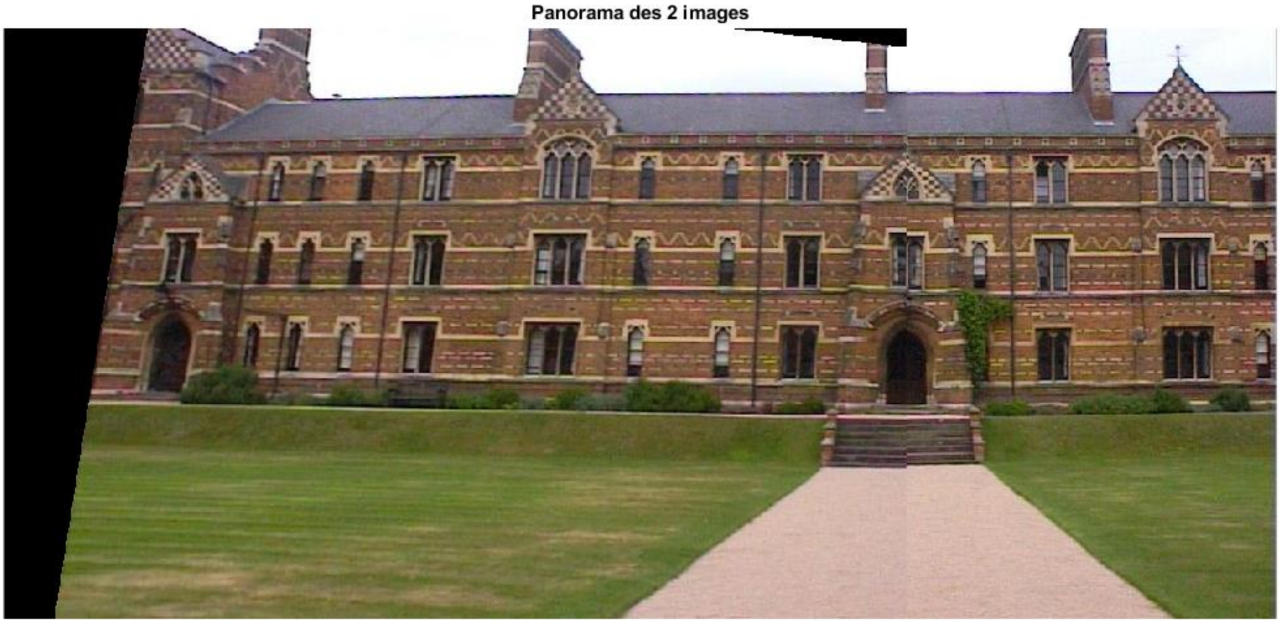
Image 2

Nous résolvons au sens des moindres carré l'équation décrite dans la partie théorie pour retrouver H et ainsi obtenir l'image 1 dans le repère image de l'image 2 :

Image 1 tranformée



On observe bien que le château présent dans l'image 1 est cette fois-ci droit comme dans l'image 2. Les deux images étant dans le même repère image il nous suffit de les combiner pour obtenir le panorama des deux images :

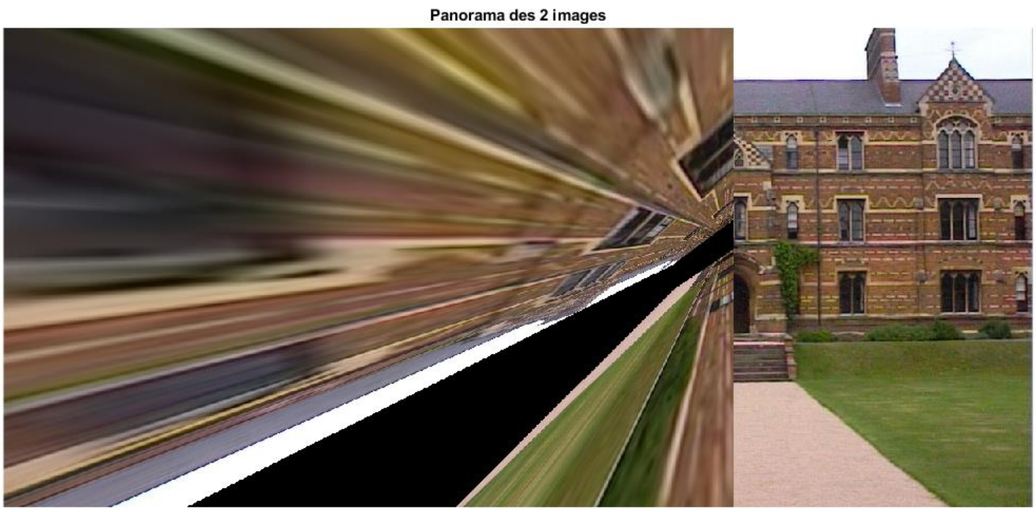


Cependant le résultat obtenu n'est pas parfait et la jointure entre des deux images est encore facilement visible. Cela est du au fait que la résolution au sens des moindres carré n'est pas une solution parfaite mais approché. Afin d'améliorer le résultat nous allons utiliser plus de point pixel entre les deux images que seulement 4 afin de réduire l'erreur résultante.

Pour les obtenir nous utilisons un algorithme qui détecte les points qui correspondent entre les deux images.

Algorithme RANSAC

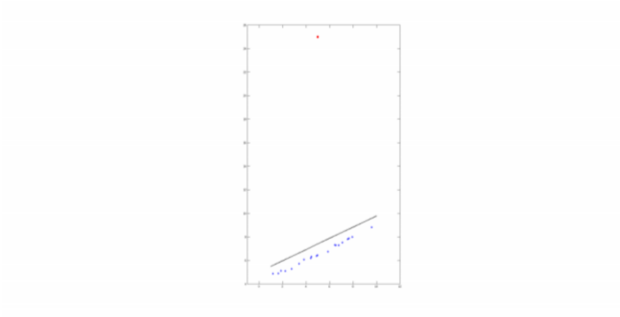
Si on essaie d'utiliser les points donnés par l'algorithme de détection on remarque que le résultat n'est pas correct comme par exemple :



En effet cela est dû au fait que parmi les points donnés par l'algorithme de détection certain points ne sont pas correct et l'erreur généré par ces points peut être très important, d'autant plus avec une approximation en moindre carré.

En effet la solution en termes des moindres carrés est très sensible aux points isolés comme nous pouvons le voir dans l'exemple si dessous :

With outliers



L^2

Dans lequel un seul point éloigné des autres déplace de manière significative la droite résultante

Pour pallier à ce problème l'algorithme RANSAC est une solution.

Le principe de cet algorithme est de choisir un petit nombre de point (16 dans l'exemple) parmi les points fournis par l'algorithme de détection. Nous réalisons la résolution avec les points choisis en évaluant la solution obtenue en comparant le nombre de pixel commun avec l'image 2. Nous répétons l'expérience en sauvegardant la meilleure solution.

De plus en notant

- w la probabilité d'obtenir un bon point parmi les points donnés par l'algorithme de détection
- m le nombre de point que l'on choisit parmi les n points donné l'algorithme de détection
- k le nombre de répétition de l'expérience on obtient la probabilité d'obtenir au moins une fois que des bons points au bout des k itérations est de :

$$p = 1 - (1 - w^m)^k$$

En voulant par exemple une probabilité de p on peut estimer k tel que

$$k = \frac{ln(1-p)}{ln(1-w^m)}$$

À la fin de l'algorithme nous avons le meilleur H et nous réalisons les transformée dans le repère image de l'image 2 tout comme précédemment :

Image 1

Image 1 tranformée

Image 2

Image 2 tranformée

Ce qui nous permet d'obtenir les panoramas suivants :

Panorama des 2 images par concatenation

Panorama des 2 images par moyenne

On observe qu’avec 16 bons points choisis la solution obtenue est visuellement parfaitement aligné à l’image 2.

De plus la création du panorama par moyenne ou par concaténation peuvent avoir leur avantage et inconvénient

Bonus

Pour la création d’un panorama avec plus de 3 images il faut bien projeter chaque image dans le même repère image comme par exemple le repère image de l’image 2. Pour cela il faut des points de référence entre l’image 2 et chacune des images. Une fois que l’on a toutes les images dans le même repère image il suffit de les combiner pour obtenir le panorama.

Conclusion

La création d’un panorama est plutôt un problème complexe, en effet il s’agit d’une résolution de système pour pouvoir retrouver l’homographie pour obtenir les points pixel d’une image 1 en 3D dans le repère camera de l’image 2.

La résolution de ce problème nécessite cependant d’avoir auparavant des points correspondants entre les images qui sont fiable car des mauvais peuvent très fortement fausser le résultat final.

L’algorithme RANSAC s’est révélé être une méthode très facilement réalisable et suffisante mais qui pourrai peut-être nécessiter un grand nombre de répétition pour des donnés qui possèdent peut de fiabilité ce qui pourrai devenir couteux.