

**Université Paris-Saclay**

Projet de fin de semestre

# Alignement d'un système multi-caméras pour l'étude des espaces congestionnés

Mohammed CHGHAF

Zhuzhi FAN

Encadré par M. Emanuel ALDEA

M2 Systèmes Embarqués et Traitement de l'Information

16 mars 2019



## TABLE DES MATIERES

1 Contexte et Problématiques .....	2
2 Travaux réalisés .....	2
2.1 Calibration des caméras .....	2
2.2 Alignement des caméras .....	3
3 Conclusion et perspectives .....	6
4 Références .....	6

## 1 CONTEXTE ET PROBLÉMATIQUES

Le laboratoire SATIE a préparé un dataset vidéo multi-caméras qui sera utilisé pour l'étude de la dynamique des piétons dans des espaces congestionnés. Des enregistrements ont été faits sur « Imperial College London », à deux reprises (avec trois et respectivement quatre caméras synchronisées à champ recouvrant). Les positions des 4 caméras sont indiquées par la Figure 1.

L'étape préliminaire à l'analyse des mouvements des piétons est l'estimation des poses des caméras et leur recalage dans un système de référence unique. Un résultat correct nous permettra de projeter des éléments visuels provenant d'une caméra dans le champ de vue de toutes les autres, et ainsi faire par la suite de la détection et du suivi de piétons de manière non supervisée. Le but de notre projet est de compléter cette étape de l'estimation des poses des caméras.

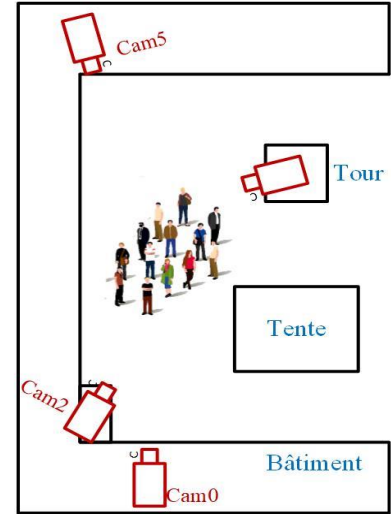


Figure 1 : Position des 4 caméras utilisées

L'objectif est résumé dans les trois points suivants :

- Obtenir l'information de positionnement et d'orientation des quatre caméras
- Pouvoir traiter des informations aux environnements compliquées : scènes homogènes et avec les objets d'intérêts répétitifs
- Réduire l'intervention humaine et la rendre pratique

## 2 TRAVAUX RÉALISÉS

### 2.1 Calibration des caméras

En raison de distorsions existants dans les photos originellement prises par des caméras, les relations spatiales entre des pixels ne peuvent pas être assurées. En conséquence, l'algorithme d'alignement ne peut pas être utilisés sur des images avec ce genre de déformation. On considère principalement deux types de déformations : distorsion radiale et distorsion tangentielle. Correspondant, il existe 5 paramètres :  $k_1, k_2, k_3, p_1, p_2$  pour décrire les deux distorsions. L'équation utilisée se trouve ci-dessous <sup>[1]</sup> :

$$\begin{pmatrix} x_p \\ y_p \end{pmatrix} = (1 + k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6) \cdot \begin{pmatrix} x_d \\ y_d \end{pmatrix} + \begin{pmatrix} 2p_1 x_d y_d + p_2 (r^2 + 2x_d^2) \\ 2p_2 x_d y_d + p_1 (r^2 + 2y_d^2) \end{pmatrix}$$

$\begin{pmatrix} x_p \\ y_p \end{pmatrix}$  : Position d'un point dans le modèle pinhole

$\begin{pmatrix} x_d \\ y_d \end{pmatrix}$  : Position déformée

$k_1, k_2, k_3$  : Coefficients de distorsion radiale

$p_1, p_2$  : Coefficients de distorsion tangentielle

$r$  : Distance entre position déformée et centre de l'image

Pour compléter la calibration, plusieurs figures de damier sont utilisées pour calculer ces 5 paramètres de chaque caméra. Ensuite, ces paramètres seront employés pour calibrer les vidéos pris à London. Le résultat de calibration des 4 caméras est présenté ci-dessous à la figure 2. On peut remarquer que la calibration des caméras a permis de retrouver les lignes naturellement droites dans les images correspondantes, ce qui était souhaité pour commencer l'étape d'alignement des caméras.

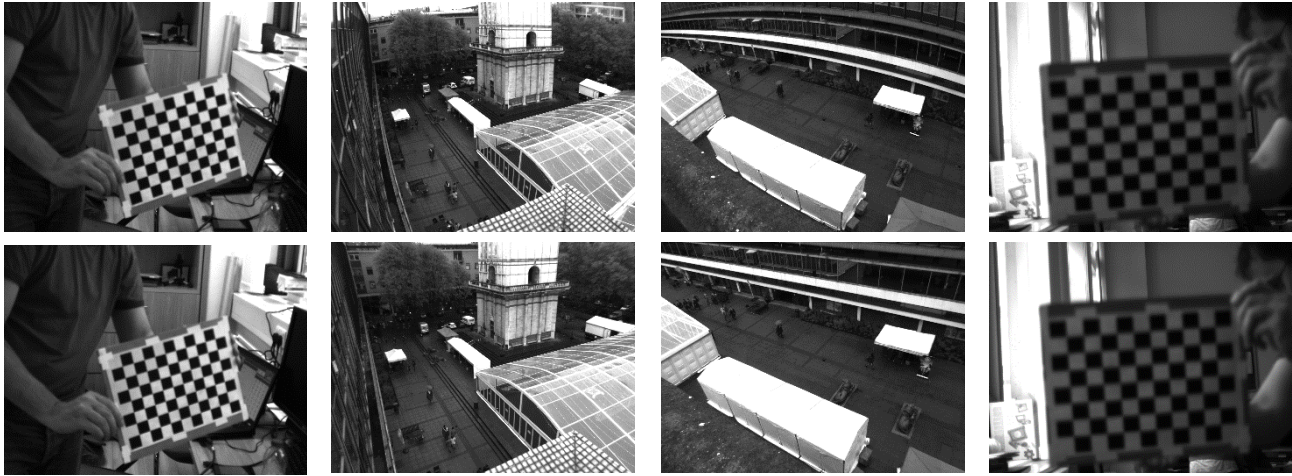


Figure 2 : Première ligne : images sont calibration, deuxième ligne : images après calibration. De gauche à droite caméra0, caméra2, caméra3 et caméra5

## 2.2 Alignement des caméras

Dans cette partie, notre but est de trouver *la matrice fondamentale*  $F$  entre plusieurs paires de caméra, qui ont les champs de vision communs assez grandes, caméra0 et caméra2, caméra2 et caméra3, caméra3 et caméra5. On calcule cette matrice en utilisant des points communs trouvé dans les paires de figures. L'équation concernant pour le calcul se trouve ci-dessous [1] :

$$p'^T F p = 0 \quad \begin{pmatrix} x_{p'} \\ y_{p'} \\ 1 \end{pmatrix}^T \cdot \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{21} \\ f_{31} & f_{32} & f_{33} \end{pmatrix} \cdot \begin{pmatrix} x_p \\ y_p \\ 1 \end{pmatrix} = 0$$

$p'$  : Vecteur de point d'objet à la première image

$p$  : Vecteur de point d'objet à la deuxième image, correspondant à  $p'$

$F$  : Matrice fondamentale

$f_{ij}$  : Élément de matrice fondamentale

Ces matrices nous permettent de connaître l'information de position relative entre des paires. Pour évaluer la performance d'alignement, la distance entre l'objet et la ligne épipolaire s'emploie pour ce but. Une illustration du principe est dans la Figure 3.

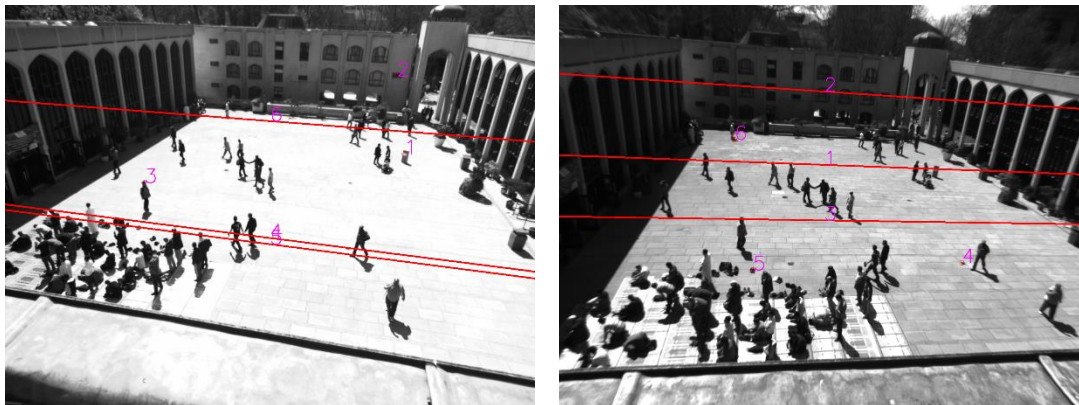


Figure 2 : Photos prises par deux caméras synchronisées. Chaque numéro dans une vue correspond au pixel ou à la ligne épipolaire correspondante dans l'autre vue.

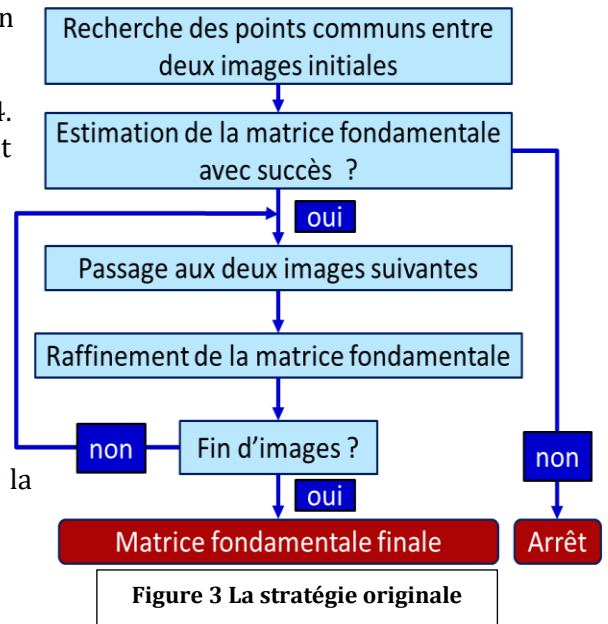
Dans cette figure (Figure 3), on peut trouver que le pixel1 vers la poubelle qu'on a choisi dans l'image1 est bien localisé dans la ligne épipolaire1 correspondant. L'erreur est moins que 3 pixels, pour les autres pixels que nous avons essayés.

Après cette étape, il faudrait relier des paires de caméras en calculant la matrice fondamentale correspondante.

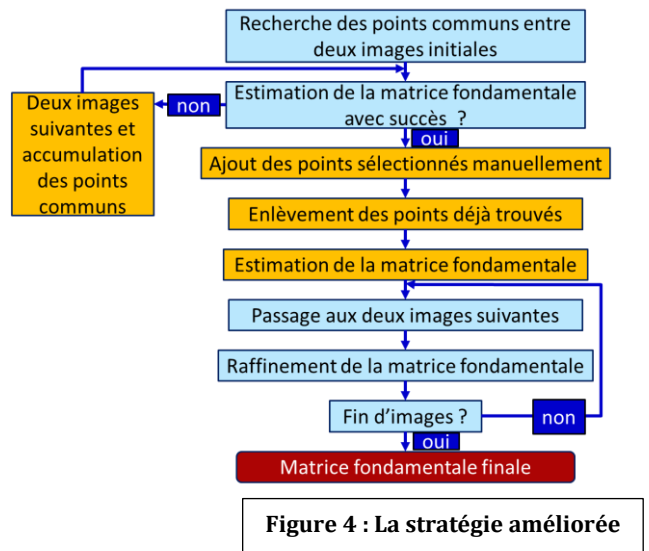
La stratégie originale <sup>[2]</sup> de notre projet est indiquée par Figure 4. Avec cette stratégie, on estime la matrice fondamentale en se basant sur des points communs trouvés entre les deux images initiales. Par la suite, on utilise des nouveaux points communs trouvés dans des images suivantes pour raffiner la matrice fondamentale.

Avec cette stratégie, il existe deux risques possibles :

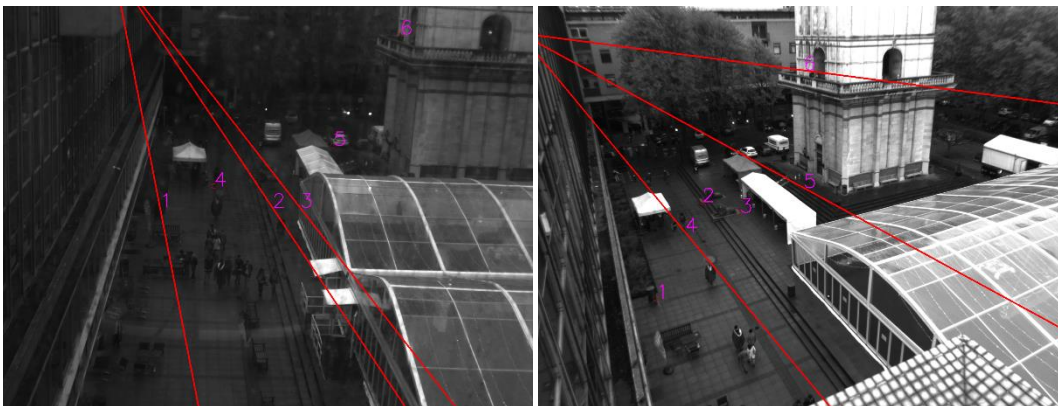
- N'ayant pas assez de points communs entre deux images initiales trouvés, l'algorithme s'arrête alors.
- Quelques points communs importants sont éliminés lorsque on réitère le processus de chercher des points communs et raffiner la matrice fondamentale.



Pour éviter ces risques, nous avons proposé d'améliorer de la stratégie originale. Avec la stratégie améliorée on n'arrête pas le programme si on ne peut pas trouver assez de points communs, par contre, on accumule des points communs jusqu'à ce qu'ils soient suffisants pour estimer la matrice fondamentale. De plus, on choisit manuellement environ 10 points communs bien répartis sur le champ de vision commun aux deux caméras. En outre, on s'assure que les points choisis manuellement ne pourront pas être éliminés par le programme. Le processus de l'algorithme est indiqué dans la Figure 5.



Le résultat de cette stratégie se trouve ci-dessous :





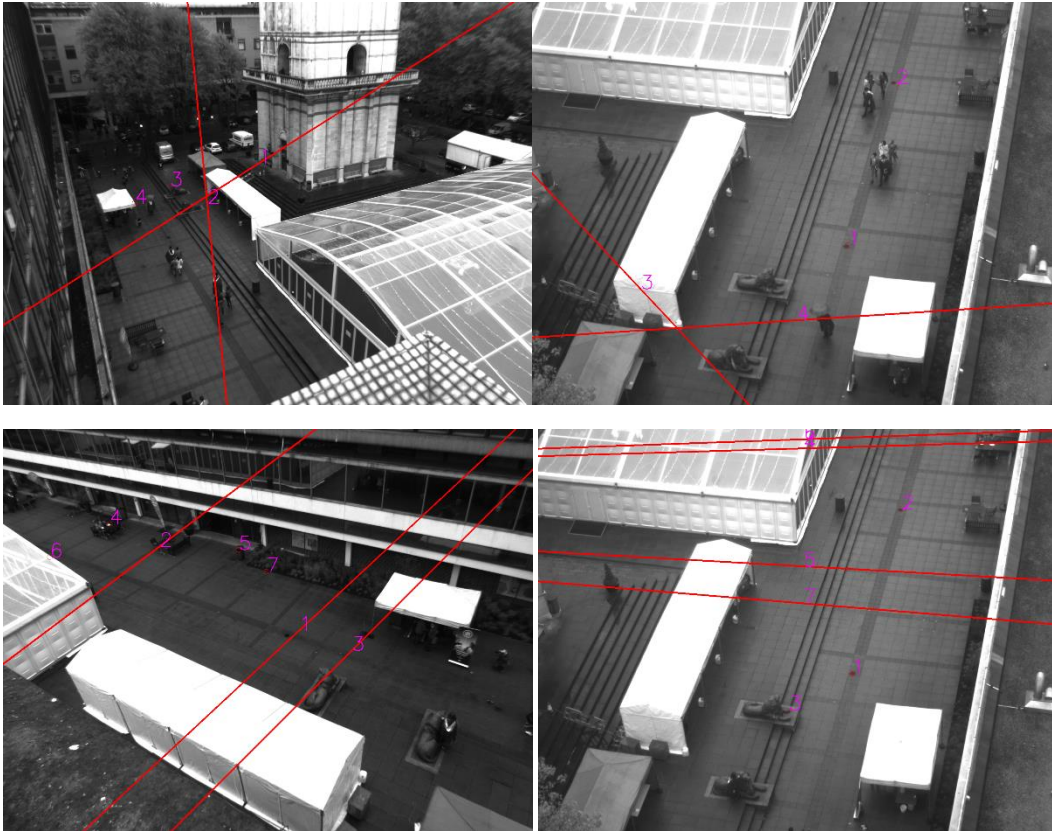


Figure 5 : De haut en bas : Résultats de l'estimation de la matrice fondamentale : entre caméra 0 et caméra2; entre caméra 2 et caméra 5; entre caméra 3 et caméra5 (Chaque paire d'images comprend des numéros qui correspondent aux pixels et à la ligne épipolaire correspondante)

La Figure 6, montre que nous avons réussi à avoir des résultats satisfaisants pour la paire (Caméra0 et Caméra2) et la paire (Caméra3 et Caméra5). Malheureusement, pour la paire (Caméra2 et Caméra5), Les résultats ne sont pas satisfaisants.

Pour mieux comprendre les causes des résultats insatisfaisants de la dernière paire, on regarde les points communs que l'algorithme retrouve pour estimer la matrice fondamentale et qui sont affichés dans la Figure 7.



Figure 6 : Caméra 2 : figure à gauche; caméra 5 : figure à droite, les points dans les deux images dont le numéro sont le même sont considéré points communs des deux images

En regardant Figure 7, nous pouvons conclure que l'imprécision de l'estimation est due aux causes suivantes :

- Le champ de vision commun à la caméra2 et la caméra5 n'est pas assez large
- Deux caméras quasiment antiparallèles, causant la difficulté de chercher des points communs
- Beaucoup de descripteurs similaires sont confondus

Puisque nous n'avons pas assez de temps pour améliorer la matrice fondamentale de la paire de (Caméra2 et Caméra5), nous ne pouvons pas continuer vers l'étape suivante consistant à relier les différentes paires de caméras.

### 3 CONCLUSION ET PERSPECTIVES

Dans ce projet, nous avons complété la calibration des 4 caméras et amélioré la stratégie d'estimation de la matrice fondamentale entre deux paires de caméras. Avec cette stratégie, nous avons estimé cette matrice des paires de 02, 25 et 35. En raison de difficulté, cités précédemment nous n'avons pas pu améliorer la matrice fondamentale entre 2 et 5, ce qui nous empêche de continuer l'étape suivante qui est d'aligner les différentes paires de caméras. Nous proposons de résoudre ce problème en considérant trois aspects ci-dessous :

- Définir la zone aveugle de chaque paire de caméras à l'étape de l'estimation de la matrice fondamentale pour réduire la confusion entre différentes régions.
- Relaxer les seuils de détection des points communs pour avoir plus de points communs dans l'étape de l'estimation de la matrice fondamentale.
- Améliorer l'algorithme pour mieux différencier les descripteurs entre deux images.

### 4 RÉFÉRENCES

[1] KAEHLER, Adrian et BRADSKI, Gary. Learning OpenCV 3: computer vision in C++ with the OpenCV library. " O'Reilly Media, Inc.", 2016.

[2] PELLICANÒ, Nicola, ALDEA, Emanuel, et LE HÉGARAT-MASCLE, Sylvie. Robust wide baseline pose estimation from video. In : 2016 23rd International Conference on Pattern Recognition (ICPR). IEEE, 2016. p. 3820-3825.