Homographies

Patrick Hébert & Denis Laurendeau (Dernière révision : septembre 2016)

Aperçu du concept d’homographie

Les homographies sont reliées à la géométrie projective impliquant 2 caméras (ou une caméra que l'on déplace) observant une ***même scène*** à partir de ***deux points de vue différents*** en ***chevauchement partiel.***

Le ***contenu*** de la scène observée doit résider sur un ***plan*** dans le référentiel "world"

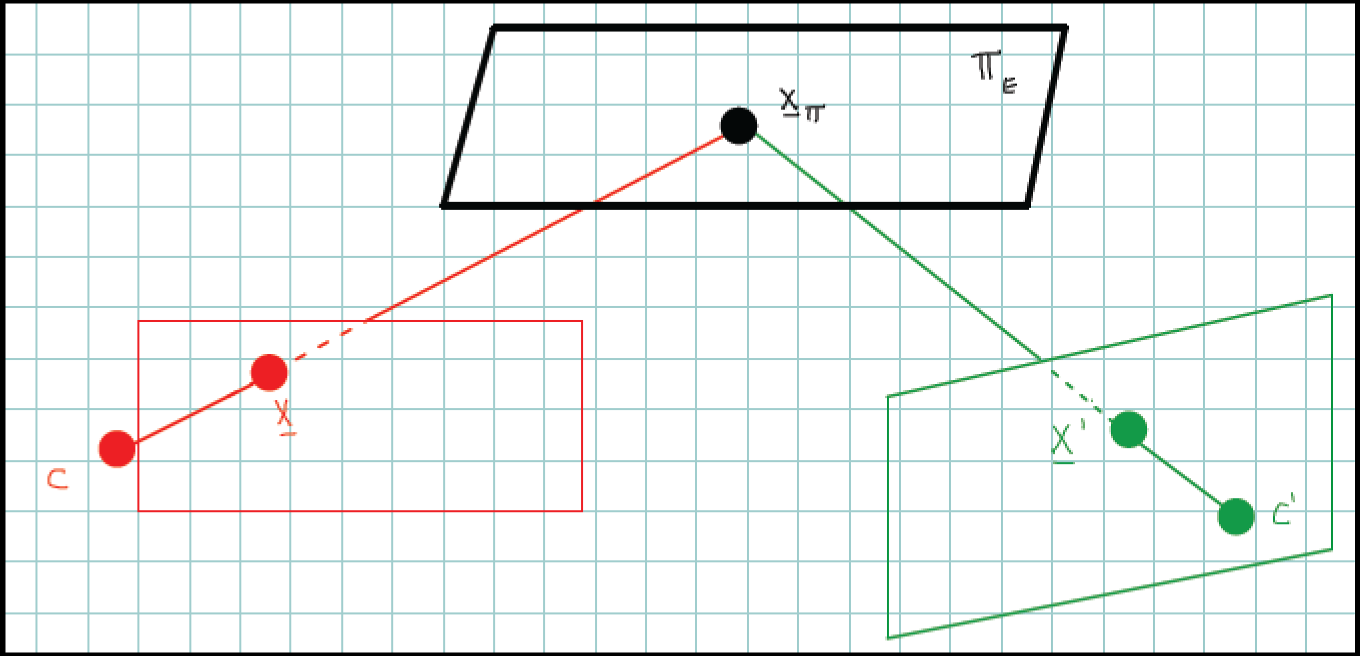


Fig. 1

Voici un exemple de scène se prêtant à l’application d’une homographie

Vue 1

Vue 2

Fig. 2

Grâce à l'homographie, on veut générer une image fusionnant les deux images de la Fig. 2 en une seule

Erreur causée par l'estimation de l'homographie.

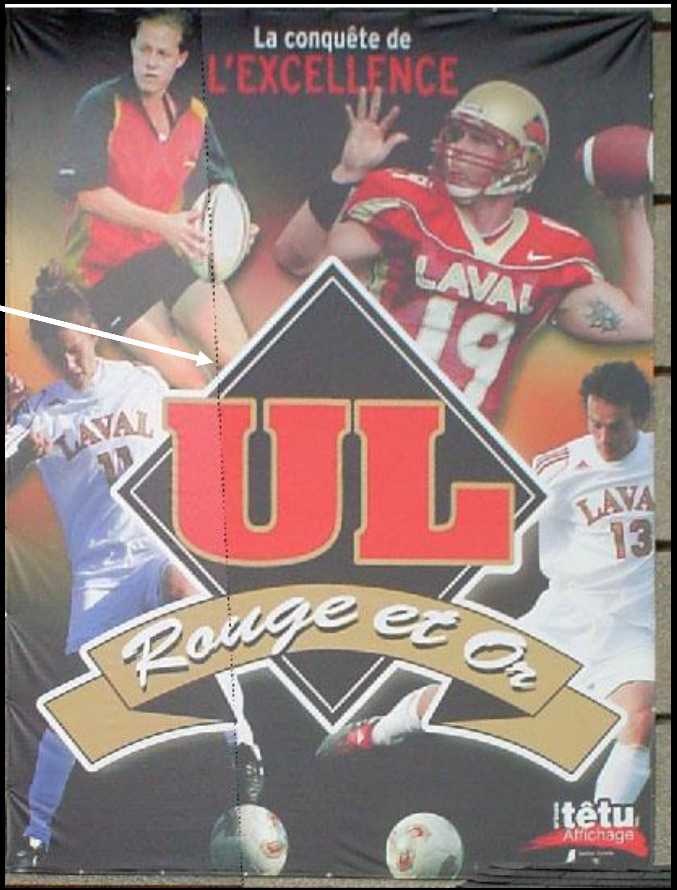


Fig. 3

Vue intégrée par l'homographie

Revenant à la Figure 1, on dit que le plan Π induit une homographie entre les deux vues.

L'homographie est une relation projective car elle n'implique que des intersections entre des droites (les projecteurs en rouge et en vert dans la Fig. 1) et ou ou des des plans dans l'espace 3D.

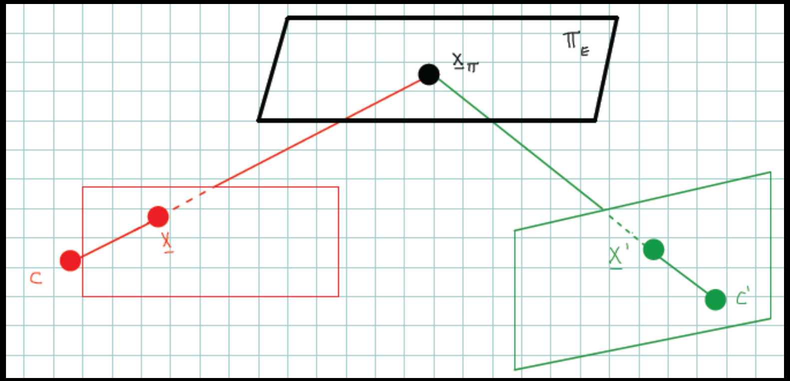


Fig. 1

L'homographie “transfère” les points-image d'une vue sur l'autre vue

Développement des fondements théoriques de l’homographie

Sur la Figure 1, nous avons une projection de perspective du point $\underline{\tilde{X}}_{\Pi}$ sur le plan image de la première vue exprimée par l’équation suivante:

$$\underline{\tilde{X}} = \underline{\underline{H}}_{1\Pi}\underline{\tilde{X}}_{\Pi} \qquad (1)$$

La seconde vue s’exprime également par une projection de perspective du point $\underline{\tilde{X}}_{\Pi}$:

$$\underline{\tilde{X}}' = \underline{\underline{H}}_{2\Pi}\underline{\tilde{X}}_{\Pi} \qquad (2)$$

De (1), on tire que:

$$\underline{\tilde{X}}_{\Pi} = \underline{\underline{H}}_{1\Pi}^{-1}\underline{\tilde{X}} \qquad (3)$$

En remplacant (3) dans (2) on a que:

$$\underline{\tilde{X}}' = \underline{\underline{H}}_{2\Pi}\underline{\underline{H}}_{1\Pi}^{-1}\underline{\tilde{X}} \qquad (4)$$

qui relie un point de l’image 1 à un point de l’image 2.

La matrice de l’équation (5) ci-dessous est appelée ***homographie***

$$\boxed{\underline{\underline{H}}_{2\Pi}\underline{\underline{H}}_{1\Pi}^{-1}} \qquad (5)$$

De manière plus compacte:

$$\underline{\tilde{X}}' = \underline{\underline{H}}\,\underline{\tilde{X}} \tag{6}$$

où:

$$\underline{\underline{H}} = \underline{\underline{H}}_{2\Pi}\underline{\underline{H}}_{1\Pi}^{-1} \tag{7}$$

est l'homographie qui permet de "transférer" des pixels de l'image 1 sur l'image 2

Dérivation analytique de l'expression de l'homographie. Arrangement de caméras simple.

Dans ce qui suit, nous développons l'expression **analytique** de l'homographie

Hypothèse 1: le plan Π dans le repère "world" est décrit par ses coordonnées dans ce repère.

Hypothèse 2: le repère de la caméra 1 est à l'origine du repère "world".

Hypothèse 3: la focale de la caméra 1 et de la caméra 2 est 1

Avec les hypothèses précédentes, les matrices caméras des sténopés des deux vues s'ecrivent comme suit:

pour la caméra 1 $$\underline{\underline{P}} = [\underline{\underline{I}} \quad \underline{0}] \tag{8}$$

pour la caméra 2 $$\underline{\underline{P}}' = [\underline{\underline{A}} \quad \underline{a}] \tag{9}$$

Le plan Π a pour sa part la forme suivante:

$$\underline{\Pi}^t \underline{\tilde{X}} = 0 \tag{10}$$

Dans (10), on choisit aussi pour simplifier:

$$\underline{\Pi} = [\underline{v}^t \quad 1]^t \tag{11}$$

ce qui signifie qu'on choisit $\Pi_4 = 1$ (la quatrième composante de $\underline{\Pi}$) pour ne pas alourdir les développements.

Le vecteur $\underline{v}$ est la normale au plan.

Dans ces conditions, la procédure pour trouver l'expression analytique de l'homographie est la suivante:

Etape 1: on effectue la projection inverse du point image de la première vue et on calcule le point d'intersection du projecteur avec le plan Π.

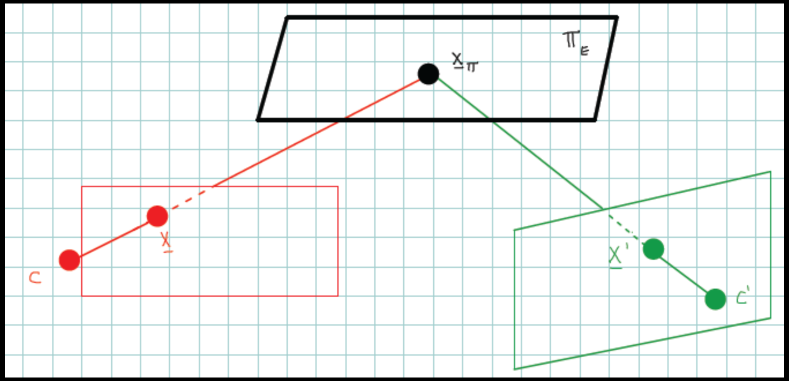


Fig. 1

Etape 2: on projette ce point d'intersection sur le plan image de la vue 2.

Pour la première vue, avec (8), la projection de perspective directe s'écrit:

$$\underline{\tilde{x}} = \underline{\underline{P}}\underline{\tilde{X}} = [\underline{\underline{I}} \quad \underline{0}]\underline{\tilde{X}} \qquad (12)$$

Dans la projection de perspective **inverse** du point image $\underline{\tilde{x}}$, un point sur le projecteur a pour coordonnées:

$$\underline{\tilde{X}} = [\underline{x}^t \quad \rho]^t \qquad (13)$$

où le paramètre ρ fixe un point sur le projecteur.

Dans (13), le point du projecteur qui intersecte le plan appartient donc au plan Π et:

$$\underline{\Pi}^t \underline{\tilde{X}} = 0 \qquad (14)$$

avec (11) et (14) on peut développer ce qui suit:

$$[\underline{v}^t \quad 1][\underline{x} \quad \rho] = 0$$

$$\underline{v}^t \underline{x} + \rho = 0$$

en isolant ρ:

$$\boxed{\rho = -\underline{v}^t \underline{x}} \qquad (15)$$

Avec (13) et (15), le point d'intersection du projecteur avec le plan Π est :

$$\underline{\tilde{X}} = [\underline{x}^t \quad -\underline{v}^t \underline{x}] \tag{16}$$

Si on projette le point de (16) sur le plan image de la deuxième vue avec (9):

$$\underline{x}' = \underline{\underline{P}}' \underline{X} = [\underline{\underline{A}} \quad \underline{a}] \underline{X} = [\underline{\underline{A}} \quad \underline{a}][\underline{x} \quad -\underline{v}^t \underline{x}]$$

on obtient:

$$\underline{\widetilde{x'}} = \underline{\underline{A}} \underline{\tilde{x}} - \underline{a} \underline{v}^t \underline{\tilde{x}} \tag{17}$$

ou encore:

$$\underline{\widetilde{x'}} = \left[\underline{\underline{A}} - \underline{a} \underline{v}^t\right] \underline{\tilde{x}} \tag{18}$$

L’expression de l’homographie est, pour le cas des caméras de (8) et (9):

$$\boxed{\underline{\underline{H}} = \underline{\underline{A}} - \underline{a}\underline{v}^t} \quad (19)$$

On remarque que la matrice renferme des informations sur les caméras ET le plan qui induit l’homographie, ce à quoi on devait intuitivement s’attendre.

Dérivation analytique de l’expression de l’homographie. Arrangement de caméras général.

L'expression (19) réfère à un arrangement de caméras simple. On peut traiter du cas d'un arrangement plus général formé d'une paire de caméras **calibrées**.

En maintenant l'hypothèse 2 sur la position de l'origine du repère "world" au centre de projection de la caméra 1, on peut maintenant écrire pour chaque caméra calibrée:

pour la caméra 1 $$\underline{\underline{P}}_E = \underline{\underline{K}}[\underline{\underline{I}} \quad \underline{0}] \tag{20}$$

pour la caméra 2 $$\underline{\underline{P}}'_E = \underline{\underline{K}}'[\underline{\underline{R}} \quad \underline{t}] \tag{21}$$

Dans (20) et (21):

Variable	Signification
$\underline{\underline{K}}$	Paramètres intrinsèques de la caméra 1
$\underline{\underline{K}}'$	Paramètres intrinsèques de la caméra 2
$\underline{\underline{R}}, \underline{t}$	Paramètres extrinsèques de la caméra 2

Posons que le plan $\underline{\Pi}_E$ s'exprime comme:

$$\Pi_E = [\underline{n}^t \quad d] \tag{22}$$

On peut écrire que, en coordonnées non-homogènes, l'équation d'un plan est:

$$\underline{n}^t \underline{X} + d = 0 \tag{23}$$

En partant du résultat (19) pour la configuration simple:

$$\underline{\underline{H}} = \underline{\underline{A}} - \underline{a}\underline{v}^t \tag{19}$$

avec, en tenant compte de (22):

$$\underline{v} = \underline{n}/d \tag{23}$$

l'homographie des caméras

$$\underline{\underline{P}} = [\underline{\underline{I}} \quad \underline{0}] \tag{24}$$

$$\underline{\underline{P}}' = [\underline{\underline{R}} \quad \underline{t}] \tag{25}$$

s’exprime comme:

$$\underline{\underline{H}} = \underline{\underline{R}} - \frac{\underline{t}\underline{n}^t}{d} \qquad (26)$$

en intégrant $\underline{\underline{K}}$ et $\underline{\underline{K}}'$ de (20) et (21) on peut développer:

$$\underline{x}' = \underline{\underline{H}}\,\underline{x}$$

$$\underline{\underline{K}}'\underline{x}' = \underline{\underline{K}}'\underline{\underline{H}}\,\underline{x}$$

$$\underline{x}'_E = \underline{\underline{K}}'\underline{\underline{H}}\,\underline{x} \qquad (27)$$

comme:

$$\underline{x} = \underline{\underline{K}}^{-1} x_E \tag{28}$$

en combinant (27) et (28):

$$\underline{x}'_E = \underline{\underline{K}}' \left[\underline{\underline{R}} \quad -\underline{t}\frac{\underline{n}^t}{d} \right] \underline{\underline{K}}^{-1} \underline{x}_E \tag{29}$$

avec **l'homographie** sous forme générale qui s'exprime comme:

$$\boxed{\underline{\underline{\boldsymbol{H}}}_{\boldsymbol{E}} = \underline{\underline{\boldsymbol{K}'}} \left[\underline{\underline{\boldsymbol{R}}} - \underline{\boldsymbol{t}}\frac{\underline{\boldsymbol{n}}^{\boldsymbol{t}}}{\boldsymbol{d}} \right] \underline{\underline{\boldsymbol{K}}}^{-\boldsymbol{1}}} \tag{30}$$

Avec ce résultat (eq. 30):

-si les caméras sont calibrées (voir chapitre précédent sur le calibrage)

-si l’équation du plan induisant l’homographie est connue

-l’homographie peut être **calculée**.

En pratique toutefois, on ne tient pas toujours à calibrer les caméras (ou il est impossible de les calibrer) et il est plus courant **d’estimer** l’homographie avec l’approche décrite dans ce qui suit.

Estimation d’une homographie

L'estimation d'une homographie se fait à partir d'un appariement de points-images provenant des deux points de vue.

L'équation de l'homographie est, comme nous l'avons vu en (6):

$$\underline{\tilde{X}}' = \underline{\underline{H}}\,\underline{\tilde{X}} \tag{6}$$

ou, de manière plus explicite:

$$\lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \underline{h}_1^t \\ \underline{h}_2^t \\ \underline{h}_3^t \end{bmatrix} \underline{\tilde{X}} \tag{31}$$

Dans l'équation (31), il y a 9 h_{ij} mais seulement 8 sont indépendants à cause du facteur d'échelle.

Si on apparie 4 points entre l'image 1 et l'image 2, on peut estimer les 8 h_{ij} indépendants car chaque point donne deux équations indépendantes. Une pour x et l'autre pour y:

$$x'_i = \frac{\underline{h}_1^t \underline{x}_i}{\underline{h}_3^t \underline{x}_i} \qquad (32)$$

$$y'_i = \frac{\underline{h}_2^t \underline{x}_i}{\underline{h}_3^t \underline{x}_i} \qquad (33)$$

qu'on peut réécrire comme suit:

$$-\underline{x}_i^t \underline{h}_1 + x'_i \underline{x}_i^t \underline{h}_3 = 0 \qquad (34)$$

$$-\underline{x}_i^t \underline{h}_2 + y'_i \underline{x}_i^t \underline{h}_3 = 0 \qquad (35)$$

En réarrangeant (34) et (35) pour prendre en compte tous les $\underline{h}_i$ et les rendre explicites:

$$-\underline{x}_i^t\underline{h}_1 + 0\underline{h}_2 + x'_i\underline{x}_i^t\underline{h}_3 = 0 \tag{36}$$

$$0\underline{h}_1 - \underline{x}_i^t\underline{h}_2 + y'_i\underline{x}_i^t\underline{h}_3 = 0 \tag{37}$$

Or, il appert que pour que l'estimation de l'homographie soit stable numériquement et soit moins sensible aux erreurs d'appariement entre points, il est préférable de construire un système d'équations surdéterminé (i.e. nombre de points appariés supérieur à 4).

Pour un nombre de points appariés N >> 4, on peut écrire le système d'équations linéaires surdéterminé suivant:

$$\boxed{\begin{bmatrix} -\underline{x}_1^t & \underline{0}_{3x1}^t & x'_1\underline{x}_1^t \\ \underline{0}_{3x1}^t & -\underline{x}_1^t & y'_1\underline{x}_1^t \\ \cdots & \cdots & \cdots \\ -\underline{x}_N^t & \underline{0}_{3x1}^t & x'_N\underline{x}_N^t \\ \underline{0}_{3x1}^t & -\underline{x}_N^t & y'_N\underline{x}_N^t \end{bmatrix}_{2Nx9} \begin{bmatrix} \underline{h}_1 \\ \underline{h}_2 \\ \underline{h}_3 \end{bmatrix}_{9x1} = \underline{0}_{2Nx1}} \quad (38)$$

qu'on peut résoudre par décomposition en valeurs singulières de la matrice de gauche de (38). La solution est la colonne de la matrice $\underline{\underline{V}}$ associée à la plus petite valeur singulière de $\underline{\underline{H}}$ dans:

$$\underline{\underline{H}} = \underline{\underline{U}}\,\underline{\underline{W}}\,\underline{\underline{V}}^t \quad (39)$$

En considérant ce qui précède, la procédure pour estimer l'homographie est la suivante.

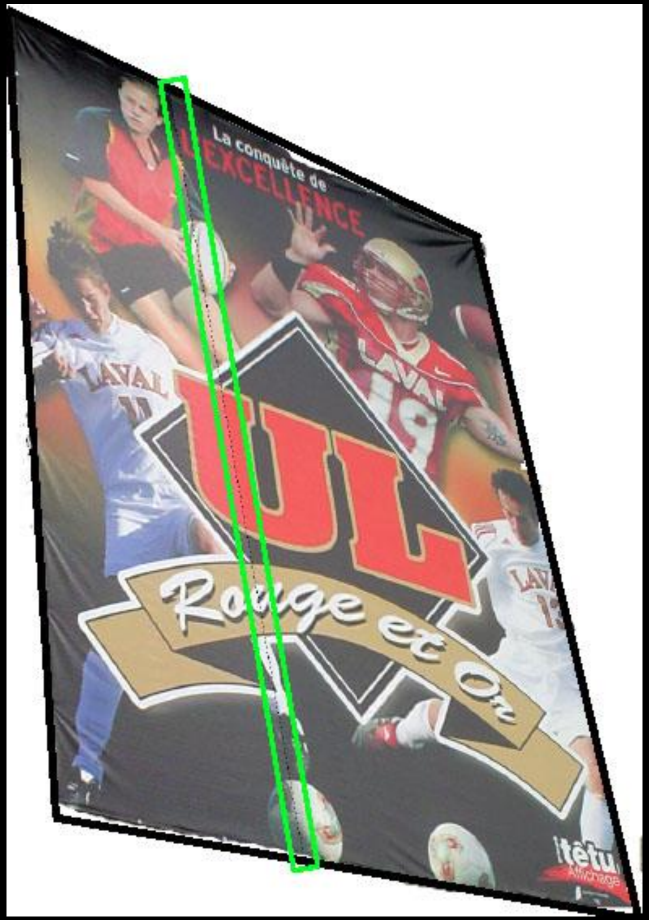
1. Apparier (manuellement ou de manière automatique) des points (avec N>>4) dans les deux images

2. Estimer l'homographie avec (38)

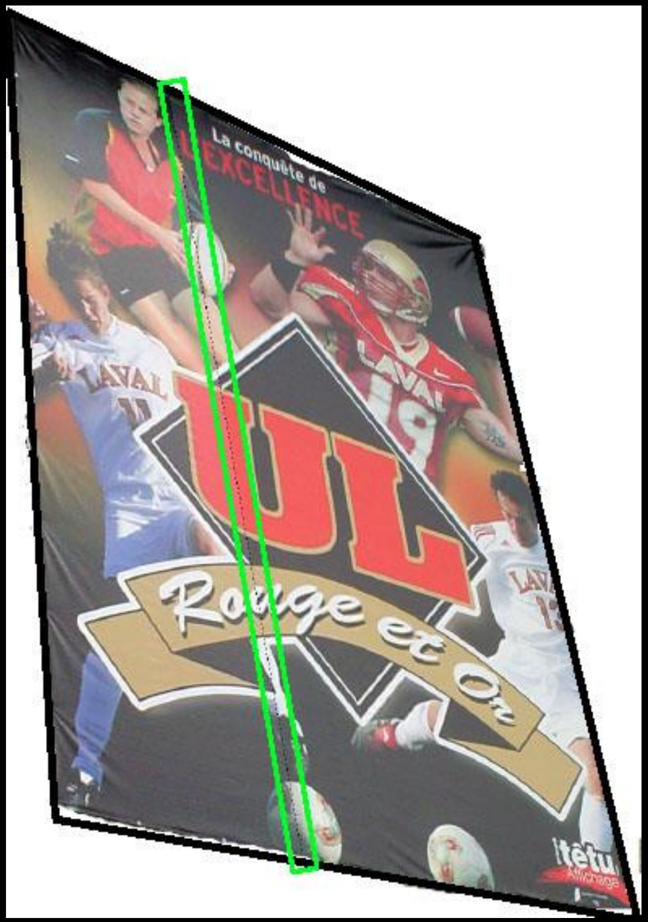
H

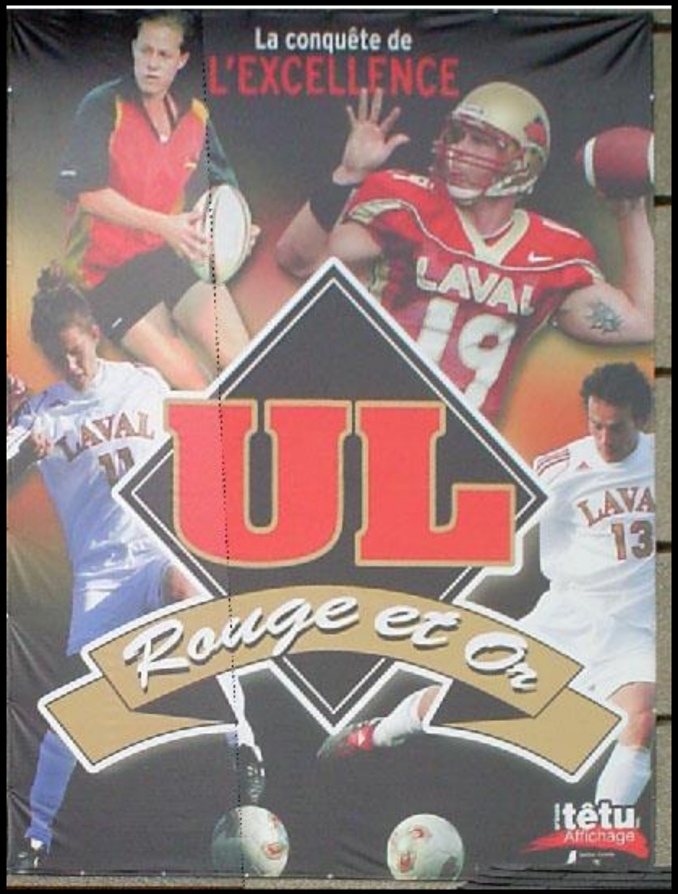
3. Redimensionner le fond de l’image 1 pour contenir le cadre de l’image 2

4. Recopier le contenu de l'image 2 dans l'image 1 grâce à l'homographie



5. Redresser l’image résultante avec une seconde homographie

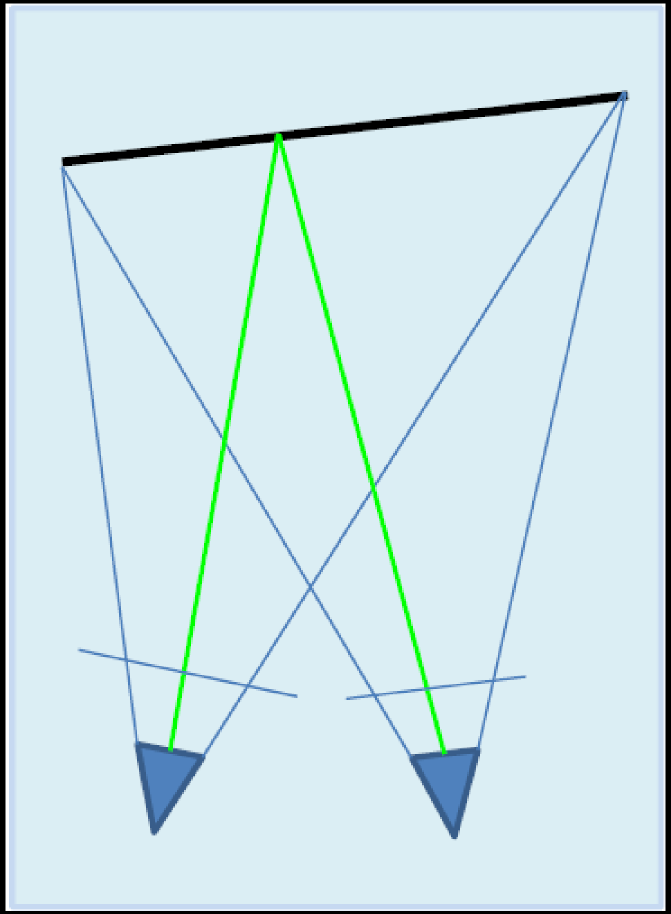


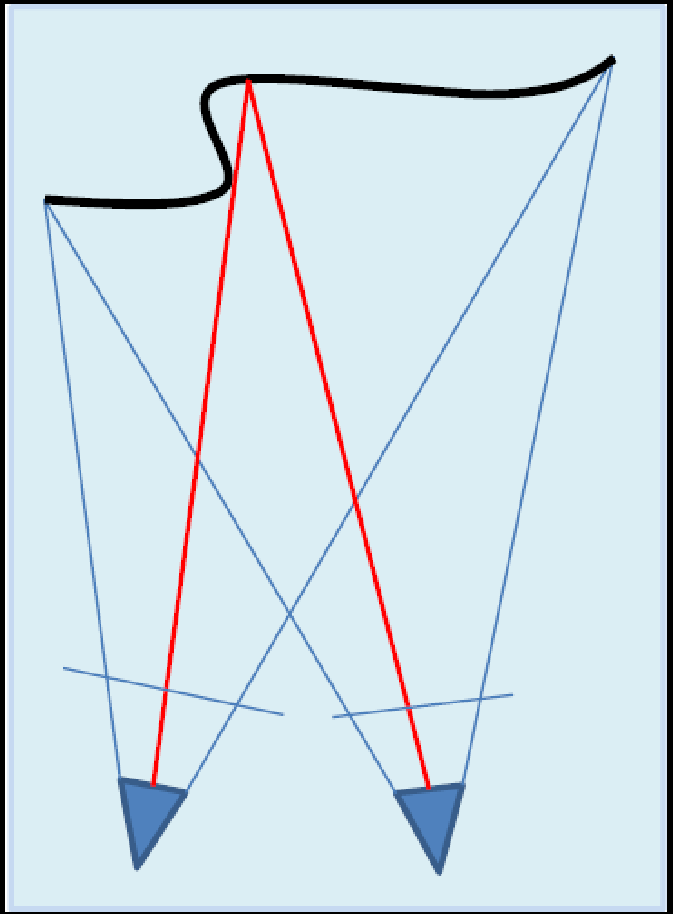


Restrictions imposées à l'homographie:

-elle est un mapping 1 à 1 et la surface induisant l'homographie doit être un plan

OK

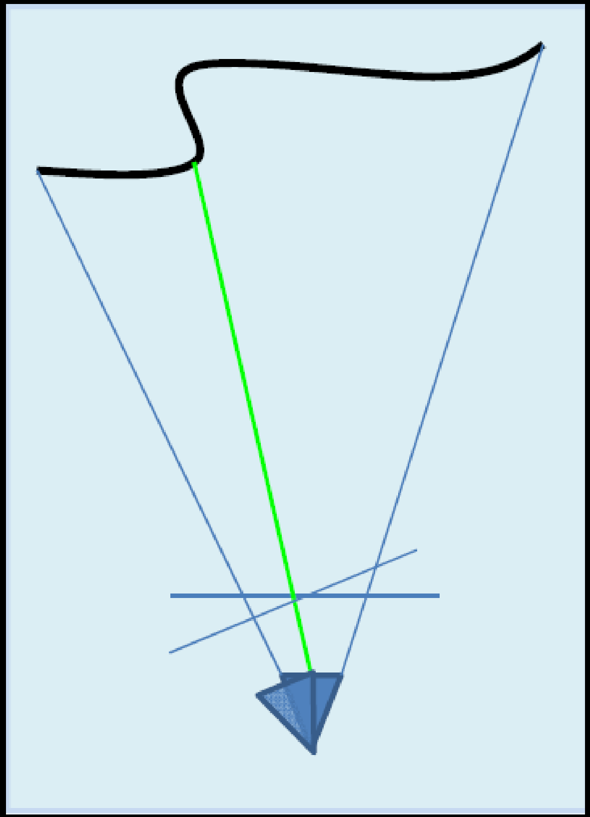




Pas OK

Restrictions imposées à l’homographie:

-une acquisition “panoramique” des images (i.e. rotation de la caméra seulement) est OK



Exemple de scène panoramique

Séquence originale

Séquence vue d'un point de vue unique synthétisé par homographie

Que faire quand plus d’un pixel tombe sur la même région de l’image?

Pour chaque point x de l’image 1:
- Regarder si il y a une valeur à Hx dans l’image 2
- Inscrire ou **fusionner** les valeurs