



---

# **MODÉLISATION MATHÉMATIQUE D'UNE CAMÉRA DE PROFONDEUR**

---

BOYER Apolline

1<sup>er</sup> août 2024

# Table des matières

1	Introduction . . . . .	2
2	Le Système Binoculaire . . . . .	2
2.1	Modélisation de la Caméra . . . . .	2
2.2	Calibration de la Caméra . . . . .	5
2.3	Vision Stéréo : Utilisation de Deux Caméras . . . . .	9
2.3.1	Principe de la Vision Stéréo . . . . .	9
2.3.2	Calcul de la Matrice de Transformation Entre les Deux Images	9
2.3.3	Géométrie Épipolaire . . . . .	12
2.4	Calcul de la Profondeur . . . . .	13
2.4.1	Principe de la Disparité . . . . .	13
2.4.2	Relation entre Disparité et Profondeur . . . . .	13
3	Time-of-Flight (ToF) Sensors . . . . .	14
3.1	Principle of ToF Sensors . . . . .	14
3.1.1	ToF Related Calculations . . . . .	15
4	Conclusion . . . . .	16

## 1 INTRODUCTION

L'objectif est de réaliser une modélisation mathématique d'une caméra de profondeur. Cette caméra de profondeur utilisera une caméra binoculaire qui permet de réaliser une vision stéréoscopique d'une image. Un capteur Time-of-Flight est également utilisé.

Tout d'abord, nous allons voir la modélisation d'un système binoculaire. Ensuite, nous verrons la modélisation d'un capteur Time-of-Flight.

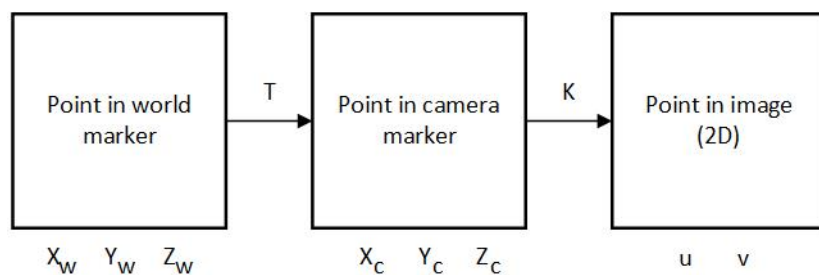
## 2 LE SYSTÈME BINOCULAIRE

La vision binoculaire, inspirée de la capacité humaine à percevoir la profondeur et la structure tridimensionnelle du monde, utilise deux caméras pour capturer des images de la même scène sous des angles légèrement différents. En exploitant la disparité, qui est la différence de position des objets dans les deux images, il est possible de calculer la distance des objets par rapport aux caméras et ainsi reconstruire une scène tridimensionnelle.

### 2.1 Modélisation de la Caméra

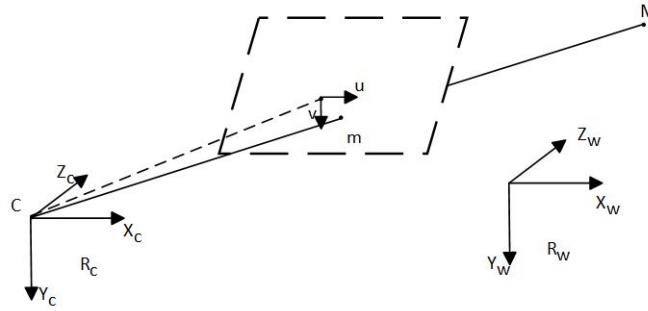
Une caméra est modélisée en utilisant le modèle à trou de serrure. Ce modèle permet la conversion d'une image 3D en une image 2D.

Cela se fait à travers trois transformations. La première transformation consiste à convertir d'un système de référence 3D au système de référence de la caméra. La deuxième transformation permet une projection sur l'image. Enfin, la troisième transformation convertit les dimensions de mm en pixels. Les deux dernières sont regroupées en une seule transformation.



**FIGURE 1** – Schéma des différentes transformations

Tout d'abord, nous prenons un point  $M$  dans ce qu'on appelle le système de coordonnées mondiales.  $M$  a pour coordonnées  $(X_W, Y_W, Z_W)$ . Le point  $C$  est la projection de  $M$   $(X_C, Y_C, Z_C)$  dans le système de référence de la caméra. Le point  $m(u, v)$  est la projection de  $C$  dans le système de référence de l'image.



**FIGURE 2** – Schéma de la transformation

Tout d'abord, nous pouvons voir la transformation de M à C comme une combinaison de rotations et de translations, le tout en coordonnées homogènes.

Les matrices de rotation autour des trois axes principaux sont définies comme suit :

**Rotation autour de l'axe X (angle  $\alpha$ ) :**

$$R_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix}$$

**Rotation autour de l'axe Y (angle  $\beta$ ) :**

$$R_y(\beta) = \begin{pmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{pmatrix}$$

**Rotation autour de l'axe Z (angle  $\gamma$ ) :**

$$R_z(\gamma) = \begin{pmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Le vecteur de translation est :

$$T(T_x, T_y, T_z) = \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix}$$

En coordonnées homogènes, les rotations et la translation peuvent être combinées en une seule matrice de transformation  $T$ .

La matrice de transformation homogène  $T$  est obtenue en combinant les matrices de rotation et le vecteur de translation :

$$T = \begin{pmatrix} R & \mathbf{t} \\ 0 & 1 \end{pmatrix} \quad (1)$$

Où  $R$  est la matrice de rotation combinée et  $\mathbf{t}$  est le vecteur de translation.

La matrice de rotation combinée peut être obtenue en multipliant les matrices de rotation individuelles (selon l'ordre souhaité des rotations) :

$$R = R_z(\gamma)R_y(\beta)R_x(\alpha)$$

Ainsi, la matrice de transformation homogène complète  $T$  est :

$$T = \begin{pmatrix} \cos\beta\cos\gamma & -\cos\beta\sin\gamma & \sin\beta & T_x \\ \sin\alpha\sin\beta\cos\gamma + \cos\alpha\sin\gamma & -\sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & -\sin\alpha\cos\beta & T_y \\ -\cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma & \cos\alpha\sin\beta\sin\gamma + \sin\alpha\cos\gamma & \cos\alpha\cos\beta & T_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2)$$

La matrice  $T$  correspond aux paramètres extrinsèques de la caméra.

Les paramètres intrinsèques d'une caméra sont utilisés pour transformer les coordonnées 3D dans le système de référence de la caméra en coordonnées 2D dans le système de référence de l'image. Ces paramètres comprennent les longueurs focales  $f_x$  et  $f_y$ , ainsi que les coordonnées du point principal  $c_x$  et  $c_y$ .

La transformation des coordonnées de la caméra  $(X_C, Y_C, Z_C)$  en coordonnées d'image  $(u, v)$  se fait en trois étapes :

1. **Projection perspective** : Cette transformation projette les coordonnées 3D sur un plan d'image.
2. **Mise à l'échelle** : Cette transformation convertit les unités des coordonnées d'image de mm en pixels.
3. **Translation** : Cette transformation ajuste les coordonnées de l'image pour tenir compte du point principal de l'image.

La matrice de projection intrinsèque  $K$  est définie comme suit :

$$K = \begin{pmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (3)$$

Ainsi, pour un point  $C$  dans le système de référence de la caméra avec des coordonnées homogènes  $(X_C, Y_C, Z_C, 1)$ , les coordonnées homogènes du point projeté  $m$  dans le système de référence de l'image sont obtenues par :

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K \begin{pmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{pmatrix} \quad (4)$$

En combinant les transformations extrinsèques et intrinsèques, la matrice de projection complète  $M$  de la caméra est :

$$M = K \cdot T = \begin{pmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} R_{11} & R_{12} & R_{13} & T_x \\ R_{21} & R_{22} & R_{23} & T_y \\ R_{31} & R_{32} & R_{33} & T_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5)$$

## 2.2 Calibration de la Caméra

L'objectif du calibrage est de déterminer les valeurs des éléments  $m_{ij}$  de la matrice  $M$  qui minimisent l'erreur de projection entre les coordonnées 3D du monde réel et les coordonnées 2D observées sur l'image. Ceci peut être accompli en utilisant des algorithmes d'optimisation tels que la méthode des moindres carrés.

Nous devons d'abord établir le système d'équations reliant les différents paramètres. Nous avons :

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = M \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix}$$

où  $(X_W, Y_W, Z_W)$  sont les coordonnées 3D du point dans le système de coordonnées du monde,  $(u, v)$  sont les coordonnées 2D de l'image en pixels,  $s$  est un facteur d'échelle, et  $M$  est la matrice de projection.

En divisant les équations par  $s$ , nous obtenons les coordonnées normalisées :

$$\begin{aligned} u &= \frac{m_{11}X_W + m_{12}Y_W + m_{13}Z_W + m_{14}}{s} \\ v &= \frac{m_{21}X_W + m_{22}Y_W + m_{23}Z_W + m_{24}}{s} \\ s &= m_{31}X_W + m_{32}Y_W + m_{33}Z_W + m_{34} \end{aligned} \quad (6)$$

En substituant  $s$  dans les équations précédentes, les coordonnées normalisées  $(u, v)$  deviennent :

$$\begin{aligned} u &= \frac{m_{11}X_W + m_{12}Y_W + m_{13}Z_W + m_{14}}{m_{31}X_W + m_{32}Y_W + m_{33}Z_W + m_{34}} \\ v &= \frac{m_{21}X_W + m_{22}Y_W + m_{23}Z_W + m_{24}}{m_{31}X_W + m_{32}Y_W + m_{33}Z_W + m_{34}} \end{aligned} \quad (7)$$

Ces équations représentent la projection d'un point du monde en un point image en utilisant la matrice  $M$  avec ses éléments  $m_{ij}$ , en considérant le facteur  $s$ .

Nous obtenons alors le système d'équations suivant :

$$\begin{cases} 0 = m_{11}X_W + m_{12}Y_W + m_{13}Z_W + m_{14} - u \cdot (m_{31}X_W + m_{32}Y_W + m_{33}Z_W + m_{34}) \\ 0 = m_{21}X_W + m_{22}Y_W + m_{23}Z_W + m_{24} - v \cdot (m_{31}X_W + m_{32}Y_W + m_{33}Z_W + m_{34}) \end{cases} \quad (8)$$

Pour calibrer la caméra, nous devons résoudre ce système d'équations pour plusieurs points  $(X_W, Y_W, Z_W)$  et leurs correspondances  $(u, v)$ . En utilisant  $N$  points, nous obtenons un système de  $2N$  équations linéaires, qui peut être exprimé sous forme matricielle pour être résolu par la méthode des moindres carrés :

$$\begin{pmatrix} X_{W1} & Y_{W1} & Z_{W1} & 1 & 0 & 0 & 0 & 0 & -u_1 X_{W1} & -u_1 Y_{W1} \\ -u_1 Z_{W1} & 0 & 0 & 0 & X_{W1} & Y_{W1} & Z_{W1} & 1 & -v_1 X_{W1} & -v_1 Y_{W1} \\ -v_1 Z_{W1} & X_{W2} & Y_{W2} & Z_{W2} & 1 & 0 & 0 & 0 & -u_2 X_{W2} & -u_2 Y_{W2} \\ -u_2 Z_{W2} & 0 & 0 & 0 & X_{W2} & Y_{W2} & Z_{W2} & 1 & -v_2 X_{W2} & -v_2 Y_{W2} \\ -v_2 Z_{W2} & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & X_{WN} & Y_{WN} & Z_{WN} & 1 & 0 & 0 & 0 & -u_N X_{WN} & -u_N Y_{WN} \\ -u_N Z_{WN} & 0 & 0 & 0 & X_{WN} & Y_{WN} & Z_{WN} & 1 & -v_N X_{WN} & -v_N Y_{WN} \\ -v_N Z_{WN} \end{pmatrix} \begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \\ m_{31} \\ m_{32} \\ m_{33} \\ m_{34} \end{pmatrix} = \begin{pmatrix} u_1 m_{34} \\ v_1 m_{34} \\ u_2 m_{34} \\ v_2 m_{34} \\ \vdots \\ u_N m_{34} \\ v_N m_{34} \end{pmatrix}$$

La méthode des moindres carrés minimise la somme des carrés des résidus, ce qui permet d'obtenir une solution optimale même en présence de bruit ou d'erreurs de mesure.

Pour réduire la complexité, nous pouvons poser  $q_{ij} \cdot m_{34} = m_{ij}$ .

Ainsi, le système d'équations devient :

$$\begin{pmatrix} X_{W1} & Y_{W1} & Z_{W1} & 1 & 0 & 0 & 0 & 0 & -u_1 X_{W1} & -u_1 Y_{W1} \\ -u_1 Z_{W1} & 0 & 0 & 0 & X_{W1} & Y_{W1} & Z_{W1} & 1 & -v_1 X_{W1} & -v_1 Y_{W1} \\ 0 & 0 & 0 & 0 & X_{W1} & Y_{W1} & Z_{W1} & 1 & -v_1 X_{W1} & -v_1 Y_{W1} \\ -v_1 Z_{W1} & X_{W2} & Y_{W2} & Z_{W2} & 1 & 0 & 0 & 0 & -u_2 X_{W2} & -u_2 Y_{W2} \\ -u_2 Z_{W2} & 0 & 0 & 0 & X_{W2} & Y_{W2} & Z_{W2} & 1 & -v_2 X_{W2} & -v_2 Y_{W2} \\ 0 & 0 & 0 & 0 & X_{W2} & Y_{W2} & Z_{W2} & 1 & -v_2 X_{W2} & -v_2 Y_{W2} \\ -v_2 Z_{W2} & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & X_{WN} & Y_{WN} & Z_{WN} & 1 & 0 & 0 & 0 & -u_N X_{WN} & -u_N Y_{WN} \\ -u_N Z_{WN} & 0 & 0 & 0 & X_{WN} & Y_{WN} & Z_{WN} & 1 & -v_N X_{WN} & -v_N Y_{WN} \\ 0 & 0 & 0 & 0 & X_{WN} & Y_{WN} & Z_{WN} & 1 & -v_N X_{WN} & -v_N Y_{WN} \\ -v_N Z_{WN} \end{pmatrix} \begin{pmatrix} q_{11} \\ q_{12} \\ q_{13} \\ q_{14} \\ q_{21} \\ q_{22} \\ q_{23} \\ q_{24} \\ q_{31} \\ q_{32} \\ q_{33} \end{pmatrix} = \begin{pmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ \vdots \\ u_N \\ v_N \end{pmatrix}$$

Nous pouvons réécrire ce problème comme suit :

$$U = K_{2N} \cdot Q \quad (9)$$

De plus, en observant les relations dans la matrice de rotation, nous pouvons déduire la relation suivante :

$$m_{31}^2 + m_{32}^2 + m_{33}^2 = 1 \quad (10)$$

D'où, nous pouvons obtenir :

$$m_{34} = \sqrt{\frac{1}{q_{31}^2 + q_{32}^2 + q_{33}^2}} \quad (11)$$

L'équation (9) est ensuite résolue par la méthode des moindres carrés. Nous obtenons le résultat :

$$Q = (K_{2N}^\top K_{2N})^{-1} K_{2N}^\top U \quad (12)$$

Une fois la matrice  $Q$  déterminée, et en utilisant l'équation (11), nous pouvons alors déterminer les éléments  $m_{ij}$ .



Ensuite, nous souhaitons calculer les paramètres intrinsèques et extrinsèques.

À partir de la relation (27), nous obtenons le système d'équations suivant :

$$\begin{aligned}
m_{11} &= f_x R_{11} + c_x R_{31} \\
m_{12} &= f_x R_{12} + c_x R_{32} \\
m_{13} &= f_x R_{13} + c_x R_{33} \\
m_{14} &= f_x T_x + c_x T_z \\
m_{21} &= f_y R_{21} + c_y R_{31} \\
m_{22} &= f_y R_{22} + c_y R_{32} \\
m_{23} &= f_y R_{23} + c_y R_{33} \\
m_{24} &= f_y T_y + c_y T_z \\
m_{31} &= R_{31} \\
m_{32} &= R_{32} \\
m_{33} &= R_{33} \\
m_{34} &= T_z
\end{aligned} \tag{13}$$

Posons  $m_i = (m_{i1}, m_{i2}, m_{i3})$ .

Les relations pour trouver tous les éléments sont :

$$\begin{aligned}
r_3 &= m_3 \\
c_x &= m_1 \cdot m_3 \\
c_y &= m_2 \cdot m_3 \\
f_x &= -\|m_1 \wedge m_3\| \\
f_y &= \|m_2 \wedge m_3\| \\
r_1 &= \frac{1}{f_x} (m_1 - c_x m_3) \\
r_2 &= \frac{1}{f_y} (m_2 - c_y m_3) \\
t_x &= \frac{1}{f_x} (m_{14} - c_x m_{34}) \\
t_y &= \frac{1}{f_y} (m_{24} - c_y m_{34}) \\
t_z &= m_{34}
\end{aligned} \tag{14}$$

## 2.3 Vision Stéréo : Utilisation de Deux Caméras

### 2.3.1 Principe de la Vision Stéréo

La vision stéréo est une technique qui utilise deux caméras placées à des positions légèrement différentes pour capturer une scène. En exploitant les différences de perspective entre les deux images obtenues, il est possible de reconstruire la structure tridimensionnelle de la scène.

Le principe de base de la vision stéréo repose sur la triangulation. En comparant les positions du même point dans les deux images, nous pouvons déterminer sa profondeur en utilisant la géométrie des triangles formés par les rayons optiques et les axes optiques des caméras.

Plus précisément, pour chaque point correspondant dans les deux images, une ligne de visée est tracée depuis le centre optique de chaque caméra vers le point. La position 3D du point est ensuite déterminée par l'intersection de ces deux lignes de visée. Ce processus est répété pour plusieurs points dans l'image afin de reconstruire la scène en trois dimensions.

### 2.3.2 Calcul de la Matrice de Transformation Entre les Deux Images

Lors de l'utilisation de deux caméras pour capturer une scène en vue stéréo, il est souvent nécessaire de calculer la matrice de transformation entre les deux images. Cette matrice, généralement appelée matrice de transformation ou matrice de déplacement, aligne les points correspondants dans les deux images pour faciliter la comparaison et la reconstruction tridimensionnelle de la scène.

La matrice de transformation est généralement calculée en utilisant des techniques d'enregistrement d'images ou de correspondance de caractéristiques. Ces techniques peuvent inclure des algorithmes tels que la correspondance de points d'intérêt, la correspondance de textures, ou des méthodes basées sur des descripteurs.

Une fois les points correspondants identifiés dans les deux images, la matrice de transformation peut être calculée à partir de ces correspondances. Cette matrice est essentiellement une transformation géométrique qui exprime le déplacement et l'orientation nécessaires pour aligner les deux images.

Soit  $A_i$  la matrice de projection qui transforme les coordonnées du cadre mondial vers le cadre de la caméra.

$$A_i = \begin{pmatrix} R_{11}^i & R_{12}^i & R_{13}^i & t_x^i \\ R_{21}^i & R_{22}^i & R_{23}^i & t_y^i \\ R_{31}^i & R_{32}^i & R_{33}^i & t_z^i \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Les matrices de projection des caméras gauche et droite sont notées  $A_L$  et  $A_R$ , respectivement. Ces matrices peuvent être définies comme suit :

$$P_{Ci} = A_i \cdot P_W \quad (15)$$

où  $P_{Ci}$  représente les coordonnées projetées dans le cadre de la caméra (gauche ou droite) et  $P_W$  représente les coordonnées du point dans le cadre mondial.

Nous pouvons alors déduire la matrice qui transforme les coordonnées du cadre de la caméra droite vers le cadre de la caméra gauche :

$$P_{CR} = A_s P_{CL} \quad (16)$$

où  $A_s = A_R \cdot A_L^{-1}$

Nous avons

$$A_s = \begin{pmatrix} R_{11} & R_{12} & R_{13} & t_x \\ R_{21} & R_{22} & R_{23} & t_y \\ R_{31} & R_{32} & R_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Nous définissons les deux points comme suit :

$$\mathbf{P}_{CR} = \begin{pmatrix} X_R \\ Y_R \\ Z_R \end{pmatrix}$$

$$\mathbf{P}_{CL} = \begin{pmatrix} X_L \\ Y_L \\ Z_L \end{pmatrix}$$

Nous obtenons alors le système d'équations suivant :

$$\begin{aligned} X_R &= R_{11} X_L + R_{12} Y_L + R_{13} Z_L + t_x \\ Y_R &= R_{21} X_L + R_{22} Y_L + R_{23} Z_L + t_y \\ Z_R &= R_{31} X_L + R_{32} Y_L + R_{33} Z_L + t_z \end{aligned} \quad (17)$$

L'objectif est d'éliminer  $Z_L$ . En posant  $x_R = \frac{X_R}{Z_R}$  et  $y_R = \frac{Y_R}{Z_R}$ , nous développons les

équations suivantes :

$$\begin{aligned} x_R &= \frac{X_R}{Z_R} = \frac{R_{11}X_L + R_{12}Y_L + R_{13}Z_L + t_x}{R_{31}X_L + R_{32}Y_L + R_{33}Z_L + t_z} \\ y_R &= \frac{Y_R}{Z_R} = \frac{R_{21}X_L + R_{22}Y_L + R_{23}Z_L + t_y}{R_{31}X_L + R_{32}Y_L + R_{33}Z_L + t_z} \end{aligned}$$

À partir de ces équations, nous pouvons exprimer  $Z_L$  comme suit :

$$\begin{aligned} Z_L &= \frac{t_x - x_R t_z}{x_R R_{31} X_L + x_R R_{32} Y_L + x_R R_{33} - R_{11} X_L - R_{12} Y_L - R_{13}} \\ Z_L &= \frac{t_y - y_R t_z}{y_R R_{31} X_L + y_R R_{32} Y_L + y_R R_{33} - R_{21} X_L - R_{22} Y_L - R_{23}} \end{aligned}$$

En égalant ces expressions, nous obtenons l'équation suivante :

$$\begin{aligned} a \cdot x_R + b \cdot y_R + c &= 0 \\ \text{où : } a &= (R_{31} t_Y - R_{21} t_Z) X_L + (R_{32} t_Y - R_{22} t_Z) Y_L + R_{33} t_Y - R_{23} t_Z \\ b &= (R_{11} t_Z - R_{31} t_X) X_L + (R_{12} t_Z - R_{32} t_X) Y_L + R_{13} t_Z - R_{33} t_X \\ c &= (R_{21} t_X - R_{11} t_Y) X_L + (R_{22} t_X - R_{12} t_Y) Y_L + R_{23} t_X - R_{13} t_Y \end{aligned} \tag{18}$$

Cette ligne permet d'estimer la correspondance entre les points dans le cadre de la caméra droite et le cadre de la caméra gauche. Elle est connue sous le nom de ligne épipolaire.

La géométrie épipolaire fournit un cadre crucial pour comprendre les relations spatiales entre les images capturées par deux caméras. Un aspect clé de cette géométrie est le concept de lignes épipolaires, qui émerge de la matrice essentielle  $E$  qui décrit les propriétés fondamentales de la scène et des caméras. De même, la matrice fondamentale  $F$  offre une représentation mathématique de la géométrie épipolaire, soulignant les relations projectives entre les points correspondants dans les images stéréo.

En effet, les lignes épipolaires sont d'une importance critique dans le processus de recherche de correspondances entre les points dans les images stéréo, une étape nécessaire pour des applications telles que la triangulation stéréo et la reconstruction tridimensionnelle.

Dans la section suivante, nous examinerons de près les calculs associés à la matrice essentielle  $E$  et à la matrice fondamentale  $F$ , ainsi que leur impact sur la détermination des lignes épipolaires. Nous explorerons également les implications pratiques de ces concepts pour la vision stéréo, en mettant en évidence les méthodes et les algorithmes utilisés pour exploiter efficacement les lignes épipolaires dans les applications pratiques.

### 2.3.3 Géométrie Épipolaire

Pour exprimer la matrice essentielle  $E$  en termes de la matrice de rotation associée à  $A_s$ , nous devons décomposer  $A_s$  en ses composants de rotation et de translation.

La matrice  $A_s$  est définie comme suit :

$$A_s = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}$$

où  $R$  est la matrice de rotation et  $t$  est le vecteur de translation.

Nous pouvons alors nous référer à l'équation de la ligne épipolaire (18). Nous pouvons exprimer les éléments  $a$ ,  $b$  et  $c$  comme suit :

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = E \cdot \begin{pmatrix} x_L \\ y_L \\ 1 \end{pmatrix} \quad (19)$$

Nous pouvons écrire  $E$  à partir de la matrice de rotation  $R$  et du vecteur de translation  $t$  :

$$E = [t]_{\times} R$$

La matrice antisymétrique  $[t]_{\times}$  est définie comme :

$$[t]_{\times} = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix}$$

En combinant ces éléments, nous obtenons :

$$E = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix} R \quad (20)$$

Nous pouvons réécrire la ligne épipolaire comme l'équation :

$$P_{CR}^t E P_{CL} = 0 \quad (21)$$

Ensuite, nous voulons calculer la matrice fondamentale  $F$ . Pour cela, nous devons revenir à la projection du cadre de la caméra vers le cadre de l'image.

$$\mathbf{p}_R = K_R \mathbf{P}_{CR}$$

$$\mathbf{p}_L = K_L \mathbf{P}_{CL}$$

En substituant les coordonnées homogènes, nous obtenons :

$$(K_R^{-1} \mathbf{p}_R)^T E (K_L^{-1} \mathbf{p}_L) = 0$$

Ce qui se développe en :

$$\mathbf{p}_R^T K_R^{-T} E K_L^{-1} \mathbf{p}_L = 0$$

Nous définissons la matrice fondamentale  $F$  en termes de la matrice essentielle  $E$  et des matrices intrinsèques :

$$F = K_R^{-T} E K_L^{-1} \quad (22)$$

L'équation épipolaire devient alors :

$$\mathbf{p}_R^T F \mathbf{p}_L = 0 \quad (23)$$

Ainsi, nous avons une relation entre les coordonnées d'image de la caméra droite et la caméra gauche.

## 2.4 Calcul de la Profondeur

Le calcul de la profondeur est une étape cruciale en vision stéréo, permettant la reconstruction de la structure tridimensionnelle d'une scène à partir d'images bidimensionnelles. En exploitant les différences de positions des points correspondants dans les deux images (disparité), il est possible de déterminer la distance de ces points par rapport aux caméras.

### 2.4.1 Principe de la Disparité

La disparité fait référence à la différence de position du même point dans les deux images capturées par les caméras stéréo. Plus précisément, si un point dans la scène est projeté sur les images gauche et droite aux coordonnées  $(x_L, y_L)$  et  $(x_R, y_R)$  respectivement, la disparité  $d$  est définie comme :

$$d = x_L - x_R \quad (24)$$

La disparité est directement liée à la profondeur du point dans la scène. Les points plus proches des caméras auront une plus grande disparité, tandis que les points plus éloignés auront une disparité plus petite.

### 2.4.2 Relation entre Disparité et Profondeur

La profondeur  $Z$  d'un point dans la scène peut être calculée en utilisant la disparité et les paramètres des caméras. Pour des caméras stéréo alignées horizontalement, la

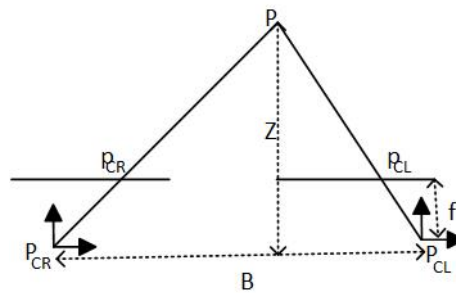
relation entre la profondeur et la disparité est donnée par :

$$Z = \frac{f \cdot B}{d} \quad (25)$$

où :

- $Z$  est la profondeur du point.
- $f$  est la distance focale des caméras.
- $B$  est la base stéréo, c'est-à-dire, la distance entre les deux caméras.
- $d$  est la disparité du point.

Le calcul de la profondeur décrit ci-dessus est valable dans le cas où les axes optiques des caméras sont parallèles et où un des axes de l'image est aligné. Cette configuration est courante dans les systèmes de vision stéréo standard, où les caméras sont alignées horizontalement et leurs plans d'image sont parallèles à la base stéréo.



**FIGURE 3** – Schéma de la vision stéréo

Il est important de noter que les écarts par rapport à ces conditions idéales peuvent introduire des erreurs dans le calcul de la profondeur. Par exemple, les distorsions optiques, les déformations d'image et les désalignements des caméras peuvent affecter la précision des mesures de disparité et introduire des erreurs dans les estimations de profondeur.

article amsmath graphicx caption geometry a4paper, margin=1in  
Depth and Time-of-Flight Sensors

### 3 TIME-OF-FLIGHT (ToF) SENSORS

#### 3.1 Principle of ToF Sensors

Time-of-Flight (ToF) sensors operate on the principle of measuring the time it takes for a light signal to travel a given distance and return to the sensor. To achieve this, the sensor emits a light signal, usually a laser beam, towards the object to be measured.

This signal is then reflected by the object and returned to the sensor. By measuring the time elapsed between the emission and reception of the signal, the sensor can determine the distance between itself and the object.

### 3.1.1 ToF Related Calculations

When the optical axes of the cameras are parallel and one of the image axes is aligned, the distance between the sensor and the object can be calculated using the following equation :

$$\text{Distance} = \frac{c \times \Delta t}{2} \quad (26)$$

In this equation :

- Distance represents the distance between the sensor and the target object.
- $c$  is the speed of light in the propagation medium (typically air).
- $\Delta t$  is the time difference between the emission and reception of the light signal.

This equation allows for the calculation of the distance between the sensor and the object using the measured time of flight and the speed of light.



## 4 CONCLUSION

Consequently, the different equations necessary for a depth camera with a binocular system and a ToF sensor are :

**Projection Matrix  $M$  :**

$$M = K \cdot T = \begin{pmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} R_{11} & R_{12} & R_{13} & T_x \\ R_{21} & R_{22} & R_{23} & T_y \\ R_{31} & R_{32} & R_{33} & T_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (27)$$

**Least Squares Result :**

$$Q = (K_{2N}^T \cdot K_{2N})^{-1} \cdot K_{2N}^T \cdot U \quad (28)$$

**Relationships to Find All Elements :**

$$\begin{aligned} r_3 &= m_3 \\ c_x &= m_1 \cdot m_3 \\ c_y &= m_2 \cdot m_3 \\ f_x &= -\|m_1 \wedge m_3\| \\ f_y &= \|m_2 \wedge m_3\| \\ r_1 &= \frac{1}{f_x} (m_1 - c_x m_3) \\ r_2 &= \frac{1}{f_y} (m_2 - c_y m_3) \\ t_x &= \frac{1}{f_x} (m_{14} - c_x m_{34}) \\ t_y &= \frac{1}{f_y} (m_{24} - c_y m_{34}) \\ t_z &= m_{34} \end{aligned} \quad (29)$$

**Essential Matrix :**

$$E = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix} R \quad (30)$$

**Fundamental Matrix  $F$  :**

$$F = K_R^{-T} E K_L^{-1} \quad (31)$$

**Epipolar Line Equation :**

$$P_{CR}^T E P_{CL} = 0 \quad (32)$$

$$\mathbf{p}_R^T F \mathbf{p}_L = 0 \quad (33)$$

**Depth Equation :**

$$Z = \frac{f \cdot B}{d} \quad (34)$$

**ToF Sensor Equation :**

$$\text{Distance} = \frac{c \times \Delta t}{2} \quad (35)$$