Lecture 9
Projective, affine and Euclidean reconstructions from two views

Olivier Faugeras

Geometric viewpoint

- Suppose we have two views of a scene and we have estimated the fundamental matrix.
- Given five point correspondences (m_i, m_i') $i = 1, \ldots, 5$ it is possible to verify from the images themselves that the 3D points M_i form a projective basis of $\mathbb{P}^3$.
- For any further correspondence (m, m'), it is then possible to compute the projective coordinates of M in the projective basis $(M_1, \ldots M_5)$.
- As a consequece, we have a 3D reconstruction of the scene which is defined up to an *unknown* projective transformation: for example the transformation that brings the canonical system of coordinates $\mathcal{R}_C$ of the first camera onto the projective basis $(M_1, \ldots M_5)$.

Algebraic viewpoint

- We compute pairs of perspective projection matrices $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{P}}'$ for the two views.
- We introduce the idea of a *canonical* representation: given a pair of perspective projection matrices $(\boldsymbol{\mathcal{P}}, \boldsymbol{\mathcal{P}}')$, the projective group $PLG(\mathbb{R}^4)$ acts upon this pair by multiplication on the right.
- Consider the *orbit* of a pair $(\boldsymbol{\mathcal{P}}, \boldsymbol{\mathcal{P}}')$, i.e. $(\boldsymbol{\mathcal{P}}_1, \boldsymbol{\mathcal{P}}'_1)$ is in the orbit if and only if there exists $H \in PLG(\mathbb{R}^4)$ represented by the 4×4 matrix $\mathbf{H}$ such that:

$$\boldsymbol{\mathcal{P}} \simeq \boldsymbol{\mathcal{P}}_1 \mathbf{H} \quad \boldsymbol{\mathcal{P}}' \simeq \boldsymbol{\mathcal{P}}'_1 \mathbf{H}$$

- A canonical representation corresponds to the choice of a representative of each orbit, e.g. the simplest in some sense. We choose in each orbit the pair $(\boldsymbol{\mathcal{P}}, \boldsymbol{\mathcal{P}}')$ such that:

$$\boldsymbol{\mathcal{P}} \simeq [\mathbf{I}_3\, \mathbf{0}]$$

Algebraic viewpoint (*continued*)

- This fixes 11 degrees of freedom for $\mathbf{H}$ which has 15 degrees of freedom. Therefore we expect to find a four-parameter family of solutions for $\boldsymbol{\mathcal{P}}'$.
- Since the epipole e' is such that:

$$\mathbf{e}' \simeq \boldsymbol{\mathcal{P}}'\mathbf{C}$$

  we have:

$$\mathbf{p}' \simeq \mathbf{e}'$$

- We determine $\mathbf{P}'$ by noticing that:

$$\mathbf{F} \simeq [\mathbf{p}']_\times \mathbf{P}'$$

  which implies that:

$$\mathbf{P}' \simeq \mathbf{H}_\Pi$$

  where Π is any plane not going through C.
- The matrix $\boldsymbol{\mathcal{P}}'$ is of the form:

$$\boldsymbol{\mathcal{P}}' \simeq [\mathbf{H}_\Pi \, \mathbf{e}']$$

Practical reconstruction

- Given $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{P}}'$, how do we compute the projective coordinates of M corresponding to the pixel pair (m, m')?
- The first method uses the projection equations:

$$m_1 : m_2 : m_3 = \langle \mathbf{\Gamma}, \mathbf{M} \rangle : \langle \mathbf{\Lambda}, \mathbf{M} \rangle : \langle \mathbf{\Theta}, \mathbf{M} \rangle$$

  and:

$$m'_1 : m'_2 : m'_3 = \langle \mathbf{\Gamma}', \mathbf{M} \rangle : \langle \mathbf{\Lambda}', \mathbf{M} \rangle : \langle \mathbf{\Theta}', \mathbf{M} \rangle$$

- Write them in linear form:

$$\begin{aligned} m_1 \langle \mathbf{\Lambda}, \mathbf{M} \rangle - m_2 \langle \mathbf{\Gamma}, \mathbf{M} \rangle = 0 \\ m_2 \langle \mathbf{\Theta}, \mathbf{M} \rangle - m_3 \langle \mathbf{\Lambda}, \mathbf{M} \rangle = 0 \\ m'_1 \langle \mathbf{\Lambda}', \mathbf{M} \rangle - m'_2 \langle \mathbf{\Gamma}', \mathbf{M} \rangle = 0 \\ m'_2 \langle \mathbf{\Theta}', \mathbf{M} \rangle - m'_3 \langle \mathbf{\Lambda}', \mathbf{M} \rangle = 0 \end{aligned}$$

- Solve the previous linear system $\mathbf{AM} = \mathbf{0}$ where $\mathbf{A}$ is 4×4 of rank 3 by SVD.

Practical reconstruction (*continued*)

- The previous method is not very good because it does not minimize any meaningful geometric quantity (i.e. $\|\mathbf{AM}\|$).
- In general it is better to minimize with respect to $\mathbf{M}$ the sum of the reprojection errors:

$$\left(\frac{m_1}{m_3} - \frac{\langle \mathbf{\Gamma}, \mathbf{M} \rangle}{\langle \mathbf{\Theta}, \mathbf{M} \rangle}\right)^2 + \left(\frac{m_2}{m_3} - \frac{\langle \mathbf{\Lambda}, \mathbf{M} \rangle}{\langle \mathbf{\Theta}, \mathbf{M} \rangle}\right)^2 +$$
$$\left(\frac{m'_1}{m'_3} - \frac{\langle \mathbf{\Gamma}', \mathbf{M} \rangle}{\langle \mathbf{\Theta}', \mathbf{M} \rangle}\right)^2 + \left(\frac{m'_2}{m'_3} - \frac{\langle \mathbf{\Lambda}', \mathbf{M} \rangle}{\langle \mathbf{\Theta}', \mathbf{M} \rangle}\right)^2$$

Canonical representation in the affine case

- The definition of a canonical representation of a pair $(\mathcal{P}, \mathcal{P}')$ is similar to the projective case, we simply replace the projective group $PLG(\mathbb{R}^4)$ by the affine group $\mathcal{GA}(\mathbb{R}^3)$. Since this group has 12 degrees of freedom and we fix 11 to ensure the special form for $\mathcal{P}$, there remains 1. We expect a one-parameter family of canonical representations.
- We still have $\mathbf{p}' \simeq \mathbf{e}'$ and furthermore:

$$\mathbf{P}'\mathbf{I}_3^{-1} \simeq \mathbf{H}_\infty$$

- The canonical representation is:

$$\mathcal{P}' \simeq [\mathbf{H}_\infty \, \mathbf{e}']$$

Nearness

- In one of the projective canonical frames, a 3D point M whose image in the first camera is m is represented by:

$$\mathbf{M} \simeq [\mathbf{m}^T \, \kappa]$$

  κ is called the nearness or planar disparity.

- Let Π be a plane not going through C, $\boldsymbol{\mathcal{P}}' \simeq [\mathbf{H}_\Pi \, \mathbf{e}']$ hence:

$$\mathbf{m}' \simeq \mathbf{H}_\Pi \mathbf{m} + \kappa \mathbf{e}'$$

- The nearness is obtained by:

$$\kappa_\Pi = \frac{(\mathbf{e}' \times \mathbf{H}_\Pi \mathbf{m})^T (\mathbf{m}' \times \mathbf{H}_\Pi \mathbf{m})}{\|\mathbf{e}' \times \mathbf{H}_\Pi \mathbf{m}\|^2}$$

Nearness (*continued*)

- Consider the Euclidean case where we use the normalized pixel coordinates:

$$\mathbf{M} \simeq \begin{bmatrix} \mathbf{m}_c \\ \frac{1}{Z} \end{bmatrix} \simeq \begin{bmatrix} xZ \\ yZ \\ Z \\ 1 \end{bmatrix}$$

- The disparity equation becomes:

$$\mathbf{m}'_c \simeq \mathbf{R}\mathbf{m}_c + \frac{1}{Z}\mathbf{t}$$

- The affine case is obtained by multiplying left by $\mathbf{A}'$:

$$\mathbf{m}' \simeq \mathbf{H}_\infty + \frac{1}{Z}\mathbf{e}'$$

- The projective case is obtained by using the relation:

$$\mathbf{H}_\Pi \simeq \mathbf{H}_\infty + \mathbf{e}'\mathbf{q}^T$$

- The disparity equation becomes:

$$\mathbf{m}' \simeq (\mathbf{H}_\Pi - \mathbf{e}'\mathbf{q}^T)\mathbf{m} + \frac{1}{Z}\mathbf{e}' = \mathbf{H}_\Pi\mathbf{m} + \kappa_\Pi\mathbf{e}'$$

- Assume that Π is represented in the canonical frame of the first camera by the vector $[\mathbf{n}^T \ -d]^T$, $d \neq 0$, then:

$$\kappa_\Pi = -\frac{d(M,\Pi)}{Zd}$$

Rectification with respect to a plane

- The idea is, given a plane Π, to rectify a stereo pair in such a way that:
  - The epipoles are at infinity in the direction of the image rows.
  - $\mathbf{H}_\Pi \simeq \mathbf{I}_3$: $M \in \Pi \Longrightarrow \kappa_\Pi = 0$.
- We apply collineations $\mathbf{H}$ and $\mathbf{H}'$ to the two image planes.
- They must satisfy:

$$\mathbf{He} \simeq \mathbf{H}'\mathbf{e}' \simeq \mathbf{e}_1 = [1, 0, 0]^T$$

  and:

$$\mathbf{H}'\mathbf{H}_\Pi\mathbf{H}^{-1} \simeq \mathbf{I}_3$$

- Since those conditions place $2+2+8 = 12$ constraints on $\mathbf{H}$ and $\mathbf{H}'$ there is a 6-dimensional family of solutions.
- We choose a solution that minimizes image distortion.

Rectification with respect to a plane (*continued*)

- Let Π be represented by $[\mathbf{n}^T\ -d]^T$, we know that:

$$\mathbf{H}_\Pi \simeq \mathbf{A}'(\mathbf{R} + \mathbf{t}\frac{\mathbf{n}^T}{d})\mathbf{A}^{-1}$$

- Apply the QR decomposition to the matrices $\mathbf{HA}$ and $\mathbf{H}'\mathbf{A}'$:

$$\mathbf{HA} = \hat{\mathbf{A}}\hat{\mathbf{R}} \quad \mathbf{H}'\mathbf{A}' = \hat{\mathbf{A}}'\hat{\mathbf{R}}'$$

- The condition $\mathbf{H}'\mathbf{e}' \simeq \mathbf{e}_1$ is equivalent to:

$$\hat{\mathbf{R}}'\mathbf{t} \simeq \mathbf{e}_1$$

- The condition $\mathbf{H}'\mathbf{H}_\Pi\mathbf{H}^{-1} \simeq \mathbf{I}_3$ is equivalent to:

$$\hat{\mathbf{R}}'\mathbf{R}\hat{\mathbf{R}} \simeq \hat{\mathbf{A}}'^{-1}\hat{\mathbf{A}} - \mathbf{e}_1\frac{\mathbf{n}^T\hat{\mathbf{R}}^T}{d}$$

Rectification with respect to a plane (*continued*)

- The right hand side is an upper triangular matrix, the left hand side is a rotation matrix hence:

$$\hat{\mathbf{R}}'\mathbf{R}\hat{\mathbf{R}} = \begin{bmatrix} \varepsilon_1 & 0 & 0 \\ 0 & \varepsilon_2 & 0 \\ 0 & 0 & \varepsilon_3 \end{bmatrix} \quad \text{with } \varepsilon_1\varepsilon_2\varepsilon_3 = 1$$

- The interpretation is that
  - H is equivalent to rotating the first camera by $\hat{\mathbf{R}}$ and changing the internal parameters to $\hat{\mathbf{A}}$.
  - H' is equivalent to rotating the second camera by $\hat{\mathbf{R}}'$ and changing the internal parameters to $\hat{\mathbf{A}}'$

Rectification with respect to a plane (*continued*)

- In terms of canonical coordinate systems:

$$\mathbf{Q}_{\hat{\mathcal{F}}_C}^{\hat{\mathcal{F}}_{C'}} = \mathbf{Q}_{\mathcal{F}_{C'}}^{\hat{\mathcal{F}}_{C'}} \mathbf{Q}_{\mathcal{F}_C}^{\mathcal{F}_{C'}} \mathbf{Q}_{\hat{\mathcal{F}}_C}^{\mathcal{F}_C} =$$

$$\begin{bmatrix} \hat{\mathbf{R}}' & \mathbf{0}_3 \\ \mathbf{0}_3^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_3^T & 1 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{R}}^T & \mathbf{0}_3 \\ \mathbf{0}_3^T & 1 \end{bmatrix} =$$

$$\begin{bmatrix} \varepsilon_1 & 0 & 0 & \alpha \\ 0 & \varepsilon_2 & 0 & 0 \\ 0 & 0 & \varepsilon_3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- This shows that the coordinate systems $\hat{\mathcal{F}}_C$ and $\hat{\mathcal{F}}_{C'}$ are parallel to each other and that the x-axis is parallel to the vector $\overrightarrow{CC'}$.

- The infinity collineation is:

$$\hat{\mathbf{H}}_\infty \simeq \hat{\mathbf{A}}' \begin{bmatrix} \varepsilon_1 & 0 & 0 \\ 0 & \varepsilon_2 & 0 \\ 0 & 0 & \varepsilon_3 \end{bmatrix} \hat{\mathbf{A}}$$

Rectification with respect to a plane (*continued*)

- We have:

$$\hat{\mathbf{H}}_{\Pi} \simeq \mathbf{I}_3$$

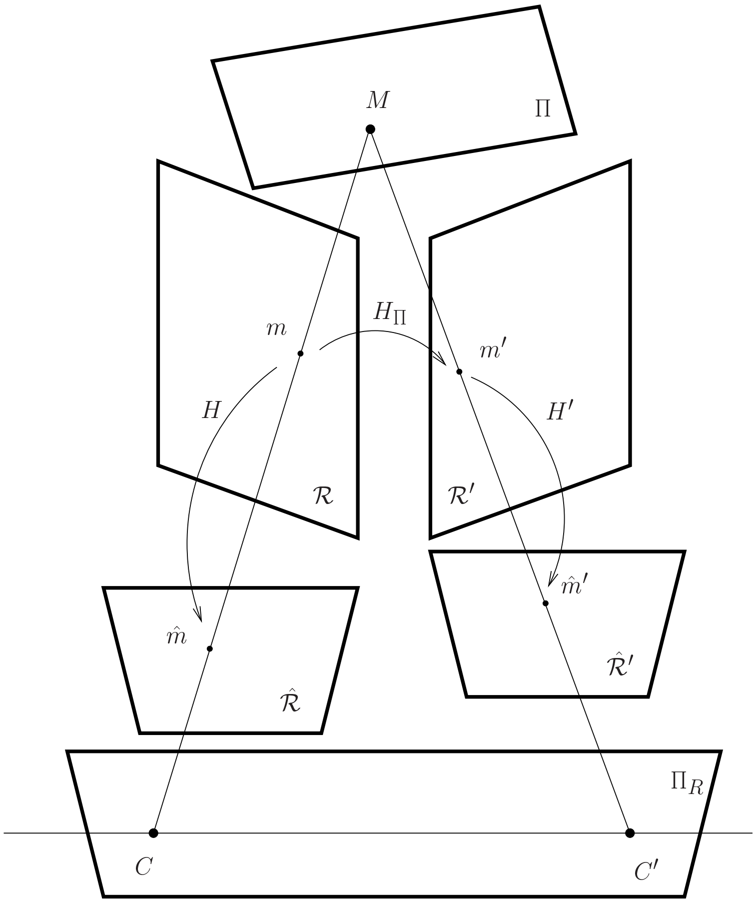


Figure 1: Rectification of a stereo pair with respect to the plane Π.

Rectification with respect to a plane (*continued*)

■ the distortion is defined as the ratio of the area of the rectangle and of the area of the white quadrangle:

$$\frac{(x_{max} - x_{min})(y_{max} - y_{min})}{\mid det(\overrightarrow{m_0 m_1}, \overrightarrow{m_0 m_2}) + det(\overrightarrow{m_0 m_2}, \overrightarrow{m_0 m_3}) \mid}$$

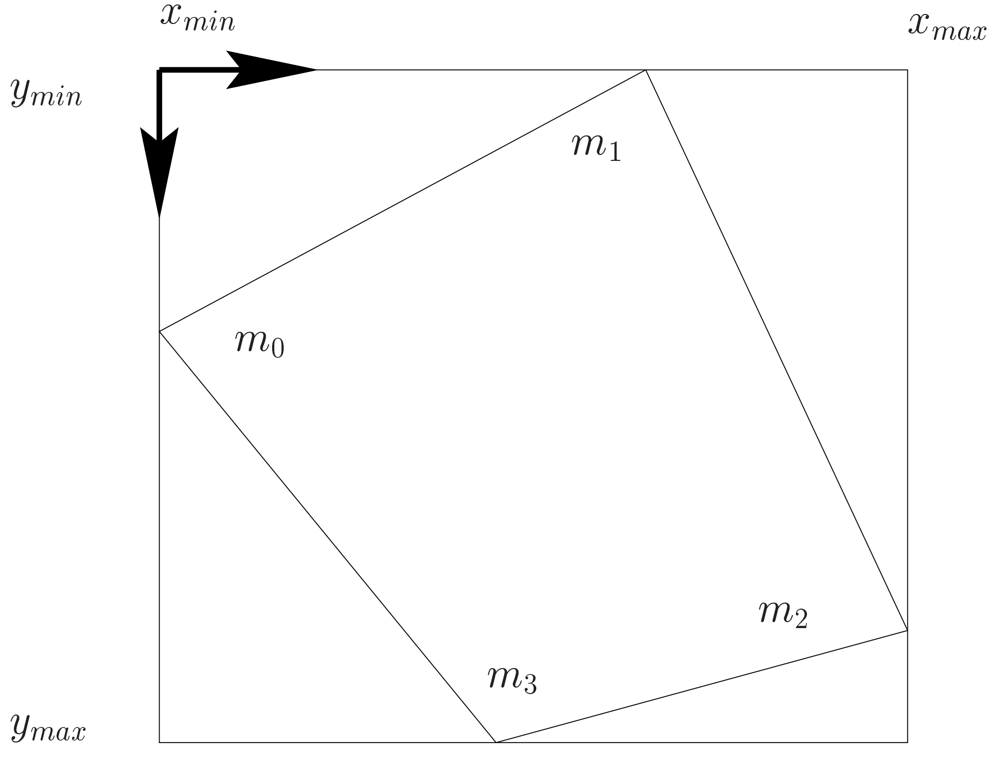


Figure 2: Definition of the distortion of the rectified image.

Application to obstacle detection

- After rectification, the disparity $x - x'$ is therefore proportional to the distance to the plane and inversely proportional to the depth.

Figure 3: A stereo pair after rectification.

Figure 4: The corresponding disparity map.

Estimation of $\mathbf{H}_\infty$

- Two methods for estimating $\mathbf{H}_\infty$.
- A priori knowledge about the scene: parallel lines, ratios of lengths.
- Controlled motion of the stereo rig, e.g. pure translation.
- Applications: affine measurements, some robotics tasks.

A priori knowledge

- Two parallel lines (L_1, L_2) provide a pair of corresponding vanishing points (m_{12}, m'_{12}) satisfying:

$$\mathbf{m}'_{12} \simeq \mathbf{H}_\infty \mathbf{m}_{12}$$

  Three such pairs are sufficient.

- Three aligned points (P, Q, R) such that the ratio $\frac{PQ}{PR}$ is known, e.g. Q is the midpoint of PR, provide a pair of corresponding vanishing points.

Translational motion

- If we can guarantee pure translation of the stereo rig, then the tracking of points or lines provides a lot of vanishing points: if (l_1, l'_1) is a correspondence *before* the translation and (l_2, l'_2) *after* the translation, then the points $\mathbf{l}_1 \times \mathbf{l}_2$ and $\mathbf{l}'_1 \times \mathbf{l}'_2$ are corresponding vanishing points.

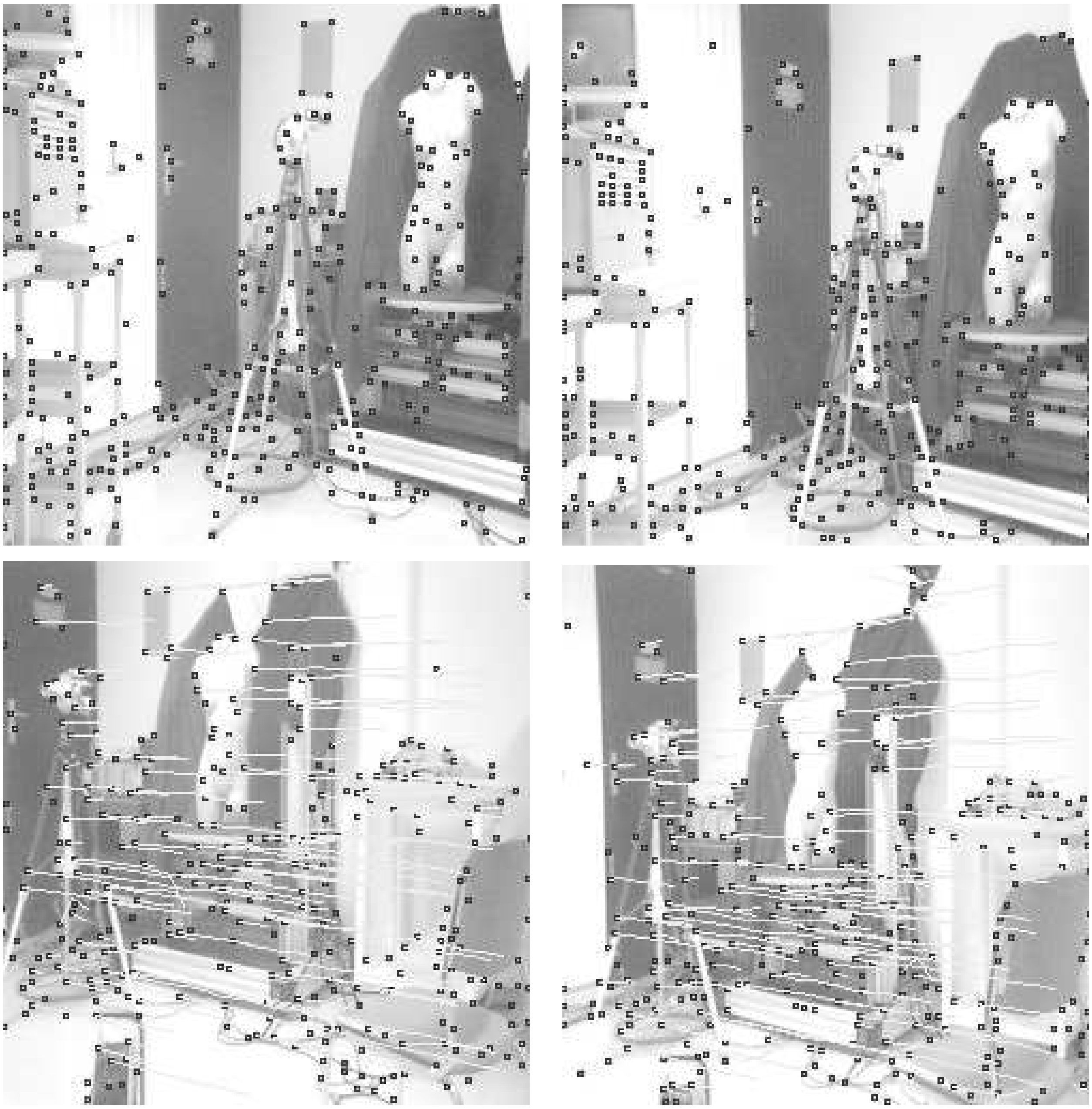
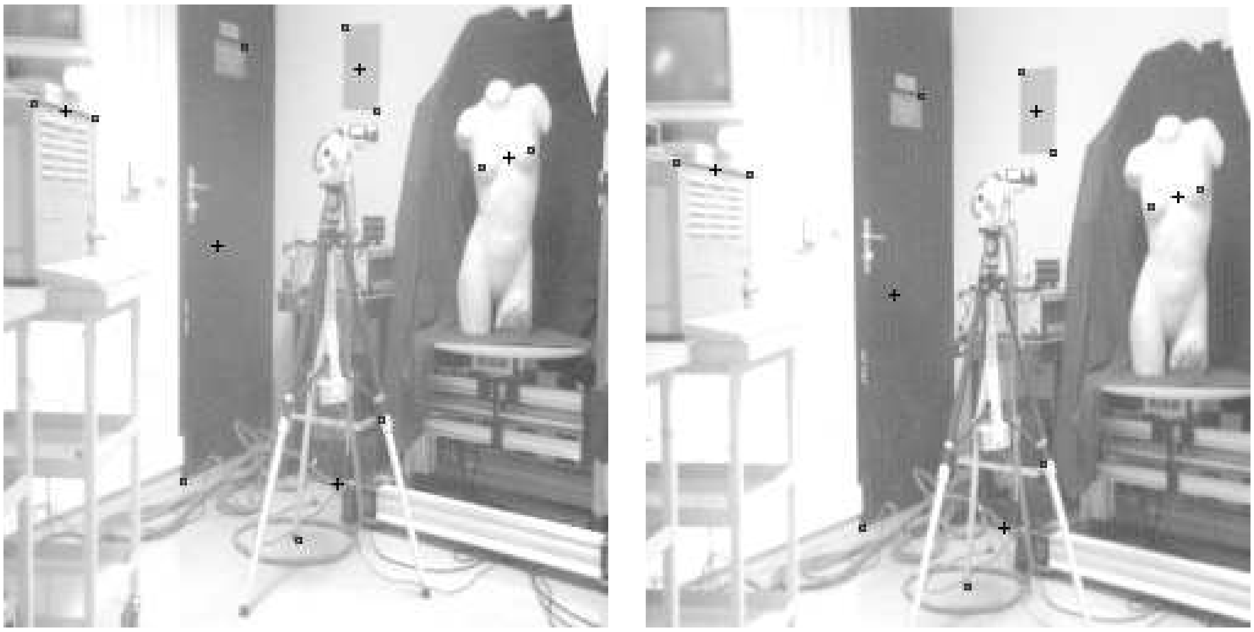


Figure 5: An example of stereo tracking.

Application: affine measurements



Application: staying in the middle of the corridor

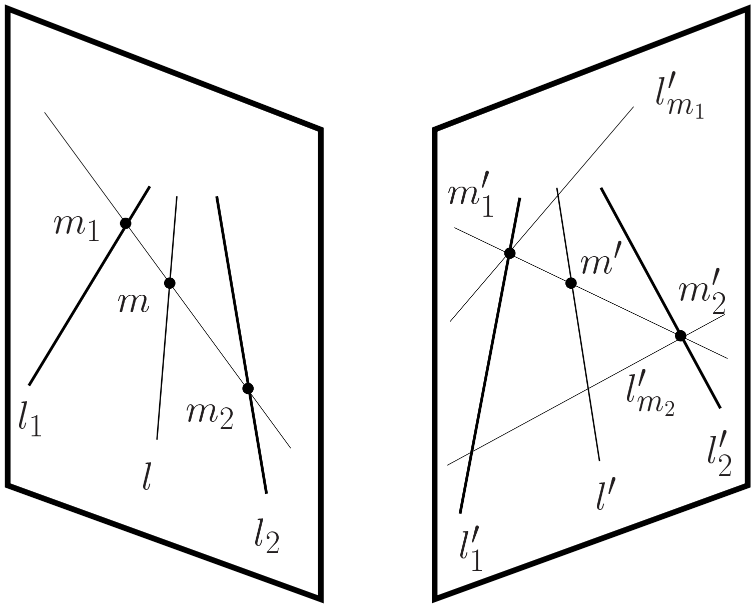


Figure 6: Theoretical determination of the middle line of a corridor.

Figure 7: Actual determination of the middle line of a corridor.

A priori information about the scene

- We assume that the affine structure has been recovered: we know $\mathbf{H}_\infty$.
- A Euclidean calibration of the rig can be obtained if we know some similitude invariants of the scene because they constrain the images of the absolute conic in the two images.
- **Angles of lines** Assume we know the angle α of two 3D lines L_1 and L_2.
- The two vanishing points are obtained from the images l_1, l_2, l'_1, l'_2 as:

$$\mathbf{m}_1 \simeq \mathbf{l}_1 \times \mathbf{H}_\infty^T \mathbf{l}'_1 \quad \mathbf{m}_2 \simeq \mathbf{l}_2 \times \mathbf{H}_\infty^T \mathbf{l}'_2$$

and:

$$\mathbf{m}'_1 \simeq \mathbf{l}'_1 \times \mathbf{H}_\infty^{-T} \mathbf{l}_1 \quad \mathbf{m}_2 \simeq \mathbf{l}'_2 \times \mathbf{H}_\infty^{-T} \mathbf{l}_2$$

- The constraints on ω and ω' are:

$$\cos\alpha = \frac{S(\mathbf{m}_1, \mathbf{m}_2)}{\sqrt{S(\mathbf{m}_1) S(\mathbf{m}_2)}} = \frac{S'(\mathbf{m}'_1, \mathbf{m}'_2)}{\sqrt{S'(\mathbf{m}'_1) S'(\mathbf{m}'_2)}}$$

A priori information about the scene (*continued*)

- **Ratio of lengths** assume we know the ratio r of the lengths of two line segments AB and CD.
- Consider the two triangles ABC and BCD, we write, for the first triangle:

$$\frac{AB}{\sin\gamma} = \frac{BC}{\sin\alpha}$$

and, for the second:

$$\frac{BC}{\sin\delta} = \frac{CD}{\sin\beta}$$

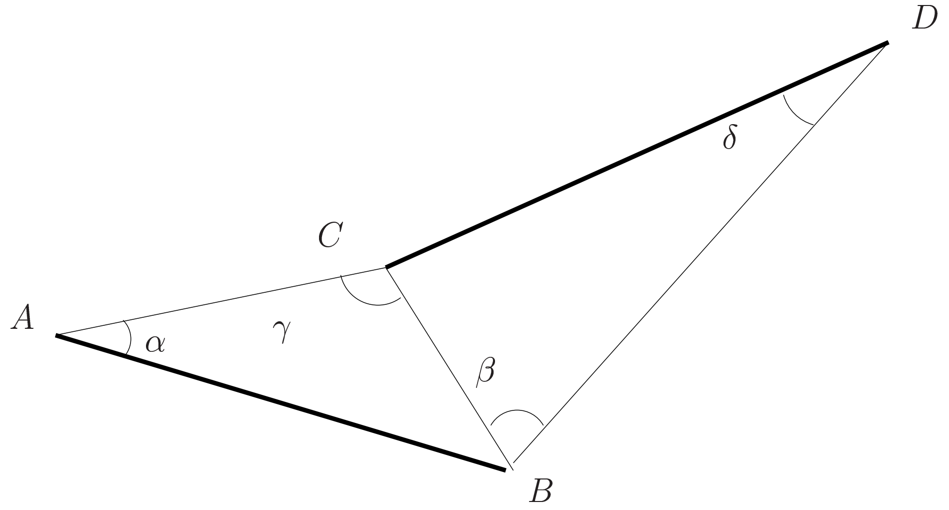


Figure 8: Some elementary euclidean geometry.

A priori information about the scene (*continued*)

- This yields the ratio $\frac{AB}{CD}$ as a function of the four angles α, β, γ and δ:

$$\frac{AB}{CD} = \frac{AB}{BC}\frac{BC}{CD} = \frac{\sin\gamma \cdot \sin\delta}{\sin\alpha \cdot \sin\beta}$$

- We construct the five vanishing points v_{ab}, v_{ac}, v_{bc}, v_{bd}, and v_{cd} of the image lines $\langle a, b\rangle$, $\langle a, c\rangle$, $\langle b, c\rangle$, $\langle b, d\rangle$ and $\langle c, d\rangle$:

$$\begin{aligned}
\mathbf{v}_{ab} &= \mathbf{l}_{ab} \times \mathbf{H}_\infty^T \mathbf{l}'_{ab}\\
\mathbf{v}_{ac} &= \mathbf{l}_{ac} \times \mathbf{H}_\infty^T \mathbf{l}'_{ac}\\
\mathbf{v}_{bc} &= \mathbf{l}_{bc} \times \mathbf{H}_\infty^T \mathbf{l}'_{bc}\\
\mathbf{v}_{bd} &= \mathbf{l}_{bd} \times \mathbf{H}_\infty^T \mathbf{l}'_{bd}\\
\mathbf{v}_{cd} &= \mathbf{l}_{cd} \times \mathbf{H}_\infty^T \mathbf{l}'_{cd}
\end{aligned}$$

- The constraint on ω is:

$$\frac{AB}{CD} = \sqrt{\frac{S(\mathbf{v}_{ab})}{S(\mathbf{v}_{cd})}} \cdot \sqrt{\frac{S(\mathbf{v}_{ac})S(\mathbf{v}_{bc}) - S^2(\mathbf{v}_{ac}, \mathbf{v}_{bc})}{S(\mathbf{v}_{ab})S(\mathbf{v}_{ac}) - S^2(\mathbf{v}_{ab}, \mathbf{v}_{ac})}} \cdot \sqrt{\frac{S(\mathbf{v}_{bd})S(\mathbf{v}_{cd}) - S^2(\mathbf{v}_{bd}, \mathbf{v}_{cd})}{S(\mathbf{v}_{bc})S(\mathbf{v}_{bd}) - S^2(\mathbf{v}_{bc}, \mathbf{v}_{bd})}}$$