21

Cheirality

When a projective reconstruction of a scene is carried out from a set of point correspondences, an important piece of information is typically ignored – if the points are visible in the images, then they must have been in front of the camera. In general, a projective reconstruction of a scene will not bear a close resemblance to the real scene when interpreted as if the coordinate frame were Euclidean. The scene is often split across the plane at infinity, as is illustrated in two dimensions by figure 21.1. It is possible to come much closer to at least an affine reconstruction of the scene by taking this simple constraint into account. The resulting reconstruction is called "quasi-affine" in that it lies part way between a projective and affine reconstruction. Scene objects are no longer split across the plane at infinity, though they may still suffer projective distortion.

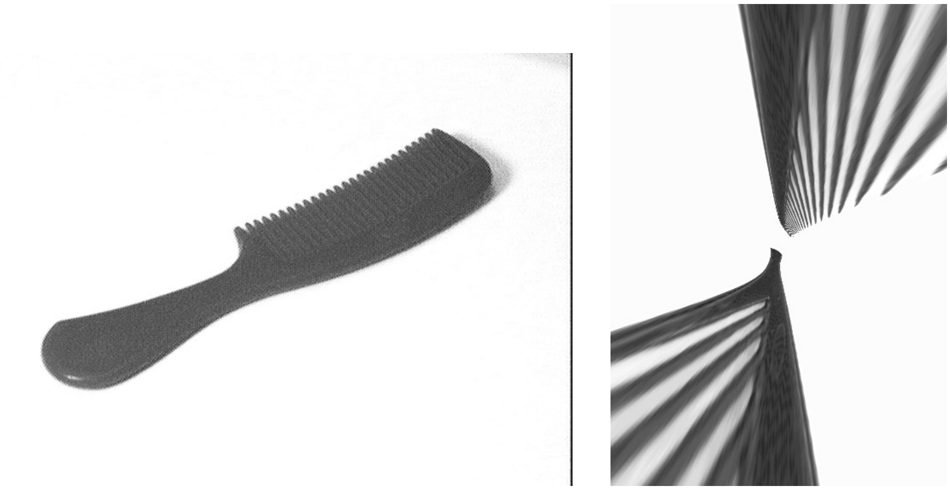
Converting a projective reconstruction to quasi-affine is extremely simple if one neglects the cameras and requires only that the scene be of the correct quasi-affine form – in fact it can be accomplished in about two lines of programming (see corollary 21.9). To handle the cameras as well requires the solution of a linear programming problem.

21.1 Quasi-affine transformations

A subset B of $\mathbb{R}^n$ is called convex if the line segment joining any two points in B also lies entirely within B. The convex hull of B, denoted $\overline{B}$, is the smallest convex set containing B. Our main concern will be with 3-dimensional point sets, so $n = 3$. We view $\mathbb{R}^3$ as being a subset of $\mathbb{P}^3$, consisting of all non-infinite points. The infinite points constitute the plane at infinity, denoted $\boldsymbol{\pi}_\infty$. Thus, $\mathbb{P}^3 = \mathbb{R}^3 \cup \boldsymbol{\pi}_\infty$. A subset of $\mathbb{P}^3$ will be called convex if and only if it is contained in $\mathbb{R}^3$ and is convex in $\mathbb{R}^3$. Hence, according to this definition, a convex set does not contain any infinite points.

Definition 21.1. Consider a point set $\{\mathbf{X}_i\} \subset \mathbb{R}^3 \subset \mathbb{P}^3$. A projective mapping $h : \mathbb{P}^3 \rightarrow \mathbb{P}^3$ is said to preserve the convex hull of the points $\{\mathbf{X}_i\}$ if

(i) $h(\mathbf{X}_i)$ is a finite point for all i, and
(ii) h maps the convex hull of points $\{\mathbf{X}_i\}$ bijectively onto the convex hull of the points $\{h(\mathbf{X}_i)\}$.



a b

Fig. 21.1. *(a) an image of a comb, and (b) the result of applying a projective transformation to the image. The projective transformation does not however preserve the convex hull of the set of points constituting the comb. In the original image, the convex hull of the comb is a finite set contained within the extent of the visible image. However, some of the points in this convex hull are mapped to infinity by the transformation.*

An example, shown in figure 21.1, may help in understanding this definition. The example deals with 2D point sets, but the principle is the same. The figure shows an image of a comb and the image resampled according to a projective mapping. The projective mapping does *not* however preserve the convex hull of the comb. Most people will agree that the resampled image is unlike any view of a comb seen by camera or human eye.

The property of preserving the convex hull of a set of points may be characterized in various different ways, as is shown by the theorem to be given shortly. In order to state this theorem, we introduce a new notation.

Notation. The symbol $\widehat{\mathbf{X}}$ denotes a homogeneous representation of a point $\mathbf{X}$ in which the last coordinate is equal to 1.

Commonly in this chapter we will be interested in the exact equalities (not equalities up to scale) between vectors representing homogeneous quantities, such as points in 3D. Thus for instance if $\mathtt{H}$ is a projective transformation, we may write $\mathtt{H}\widehat{\mathbf{X}} = \mathtt{T}'\widehat{\mathbf{X}}'$ to mean that the transformation takes point $\widehat{\mathbf{X}}$ to point $\widehat{\mathbf{X}}'$, but that the scale factor $\mathtt{T}'$ is required to make this equality exact.

Now for the theorem.

Theorem 21.2. *Consider a projective transformation $h : \mathbb{P}^3 \rightarrow \mathbb{P}^3$ and a set of points $\{\mathbf{X}_i\}$. Let $\boldsymbol{\pi}_\infty$ be the plane mapped to infinity by h. The following statements are equivalent.*

(i) *h preserves the convex hull of the points $\{\mathbf{X}_i\}$.*
(ii) *$\tilde{\mathbf{X}} \cap \boldsymbol{\pi}_\infty = \emptyset$ for any point $\tilde{\mathbf{X}}$ in the convex hull of the points $\{\mathbf{X}_i\}$.*
(iii) *Let $\mathtt{H}$ be a matrix representing the transformation h, and suppose that $\mathtt{H}\widehat{\mathbf{X}}_i = \mathtt{T}'_i\widehat{\mathbf{X}}'_i$. Then the constants $\mathtt{T}'_i$ all have the same sign.*

Proof. This will be proved by showing that $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i)$.

$(i) \Rightarrow (ii)$. If h preserves the convex hull of the points, then $h(\tilde{\mathbf{X}})$ is a finite point for any point $\tilde{\mathbf{X}}$ in the convex hull of the $\mathbf{X}_i$. So $\tilde{\mathbf{X}} \cap \boldsymbol{\pi}_\infty = \emptyset$.

$(ii) \Rightarrow (iii)$. Consider the chord joining two points $\mathbf{X}_i$ and $\mathbf{X}_j$ and suppose that $\mathtt{T}'_i$ and $\mathtt{T}'_j$ (as in part (iii) of the theorem) have opposite signs. Since $\mathtt{T}'_i$ is a continuous function of the coordinates of $\mathbf{X}_i$, there must be a point $\tilde{\mathbf{X}}$ lying on the chord from $\mathbf{X}_i$ to $\mathbf{X}_j$ for which $\mathtt{T}'$ is equal to zero. This means that $\mathtt{H}\tilde{\mathbf{X}} = (\tilde{\mathtt{X}}', \tilde{\mathtt{Y}}', \tilde{\mathtt{Z}}', 0)^\mathsf{T}$. Since $\tilde{\mathbf{X}}$ is in the convex hull of the points $\{\mathbf{X}_i\}$, this contradicts (ii).

$(iii) \Rightarrow (i)$. We assume that there exist constants $\mathtt{T}'_i$ all of the same sign such that $\mathtt{H}\widehat{\mathbf{X}}_i = \mathtt{T}'_i\widehat{\mathbf{X}}'_i$. Let S be the subset of $\mathbb{R}^n$ consisting of all points $\mathbf{X}$ satisfying the condition $\mathtt{H}\widehat{\mathbf{X}} = \mathtt{T}'\widehat{\mathbf{X}}'$ such that $\mathtt{T}'$ has the same sign as the $\mathtt{T}'_i$. The set S contains $\{\mathbf{X}_i\}$. It will be shown that S is convex. If $\mathbf{X}_i$ and $\mathbf{X}_j$ are points in S with corresponding constants $\mathtt{T}'_i$ and $\mathtt{T}'_j$, then any intermediate point $\mathbf{X}$ between $\mathbf{X}_i$ and $\mathbf{X}_j$ must have $\mathtt{T}'$ value intermediate between $\mathtt{T}'_i$ and $\mathtt{T}'_j$. To see this, consider a point $\widehat{\mathbf{X}} = \alpha\widehat{\mathbf{X}}_i + (1-\alpha)\widehat{\mathbf{X}}_j$ where $0 \le \alpha \le 1$. This point lies between $\mathbf{X}_i$ and $\mathbf{X}_j$. Denote by $\mathbf{h}_4^\mathsf{T}$ the last row of $\mathtt{H}$. Then,

$$\begin{aligned} \mathtt{T}' &= \mathbf{h}_4^\mathsf{T}\widehat{\mathbf{X}} \\ &= \mathbf{h}_4^\mathsf{T}(\alpha\widehat{\mathbf{X}}_i + (1-\alpha)\widehat{\mathbf{X}}_j) \\ &= \alpha\mathbf{h}_4^\mathsf{T}\widehat{\mathbf{X}}_i + (1-\alpha)\mathbf{h}_4^\mathsf{T}\widehat{\mathbf{X}}_j \\ &= \alpha\mathtt{T}'_i + (1-\alpha)\mathtt{T}'_j \end{aligned}$$

which lies between $\mathtt{T}'_i$ and $\mathtt{T}'_j$ as claimed. Consequently, the value of $\mathtt{T}'$ must have the same sign as $\mathtt{T}'_i$ and $\mathtt{T}'_j$, and so $\mathbf{X}$ lies in S also. This shows that S is convex.

Now, let $\tilde{S}$ be a convex subset of S. It will be shown that $h(\tilde{S})$ is also convex. This is easily seen to be true, since h maps a line segment in $\tilde{S}$ to a line segment that does not cross the plane at infinity. Thus, h maps any convex set $\tilde{S}$ such that $S \supset \tilde{S} \supset \{\mathbf{X}_i\}$ to a convex set $\tilde{S}'$ such that $S' \supset \tilde{S}' \supset \{\mathbf{X}'_i\}$. However, if $\mathtt{H}$ satisfies condition (iii), then it is easily seen that $\mathtt{H}^{-1}$ does also. From this it follows that the above correspondence between convex sets $\tilde{S}$ and $\tilde{S}'$ is bijective. Since the convex hull of $\{\mathbf{X}_i\}$ (or $\{\mathbf{X}'_i\}$ respectively) is the intersection of all such convex sets, it follows that h preserves the convex hull of the points. $\square$

The projective transformations that preserve the convex hull of a given set of points form an important class, and will be called *quasi-affine* transformations.

Definition 21.3. Let B be a subset of $\mathbb{R}^n$ and let h be a projectivity of $\mathbb{P}^n$. The projectivity h is said to be "quasi-affine" with respect to the set B if h preserves the convex hull of the set B.

It may be verified that if h is quasi-affine with respect to B, then h^{-1} is quasi-affine with respect to $h(B)$. Furthermore, if h is quasi-affine with respect to B and g is quasi-affine with respect to $h(B)$, then $g \circ h$ is quasi-affine with respect to B. Thus, quasi-affine projectivities may be composed in this fashion. Strictly speaking, however, quasi-affine projectivities with respect to a given fixed set of points do not form a group.

We will be considering sets of points $\{\mathbf{X}_i\}$ and $\{\mathbf{X}'_i\}$ that correspond via a projectivity. When we speak of the projectivity being *quasi-affine*, we will mean with respect to the set $\{\mathbf{X}_i\}$.

Two-dimensional quasi-affine mappings

Two-dimensional quasi-affine mappings arise as transformations between planar point sets in 3D and their image under a projective camera mapping, as stated formally below.

Theorem 21.4. *If B is a point set in a plane (the "object plane") in $\mathbb{R}^3$ and B lies entirely in front of a projective camera, then the mapping from the object plane to the image plane defined by the camera is quasi-affine with respect to B.*

Proof. That there is a projectivity h mapping the object plane to the image plane is well known. What is to be proved is that the projectivity is quasi-affine with respect to B. Let L be the line in which the principal plane of the camera meets the object plane. Since B lies entirely in front of the camera, L does not meet the convex hull of B. However, by definition of the principal plane $h(L) = L_\infty$, where L_∞ is the line at infinity in the image plane. Thus, one deduces that $h(\overline{B}) \cap L_\infty = \emptyset$, and hence by theorem 21.2 the transformation is quasi-affine with respect to B. $\square$

Note that if points $\mathbf{x}_i$ are visible in an image, then the corresponding object points must lie in front of the camera. Applying theorem 21.4 to a sequence of imaging operations (for instance, a picture of a picture of a picture, etc.), it follows that the original and final images in the sequence are connected by a planar projectivity which is quasi-affine with respect to any point set in the object plane visible in the final image.

Similarly, if two images are taken of a set of points $\{\mathbf{X}_i\}$ in a plane, $\{\mathbf{x}_i\}$ and $\{\mathbf{x}'_i\}$ being corresponding points in the two images, then there is a quasi-affine mapping (with respect to the $\mathbf{x}_i$) mapping each $\mathbf{x}_i$ to $\mathbf{x}'_i$, and so theorem 21.2 applies, yielding the following:

Result 21.5. *If $\{\mathbf{x}_i\}$ and $\{\mathbf{x}'_i\}$ are corresponding points in two views of a set of object points $\{\mathbf{X}_i\}$ lying in a plane, then there is a matrix $\mathtt{H}$ representing a planar projectivity such that $\mathtt{H}\hat{\mathbf{x}}_i = w_i\hat{\mathbf{x}}'_i$ and all w_i have the same sign.*

21.2 Front and back of a camera

The depth of a point $\mathbf{X} = (\mathrm{X}, \mathrm{Y}, \mathrm{Z}, \mathrm{T})^\mathsf{T}$ with respect to a camera was shown in (6.15–*p*162) to be given by

$$\mathrm{depth}(\mathbf{X}; \mathtt{P}) = \frac{\mathrm{sign}(\det \mathtt{M})w}{\mathrm{T}\|\mathbf{m}^3\|} \tag{21.1}$$

where $\mathtt{M}$ is the left hand 3×3 block of $\mathtt{P}$, $\mathbf{m}^3$ is the third row of $\mathtt{M}$, and $\mathtt{P}\mathbf{X} = w\hat{\mathbf{x}}$. This expression is not dependent on the particular homogeneous representation of $\mathbf{X}$ or $\mathtt{M}$, that is it is unchanged by multiplication by non-zero scale factors. This definition of depth is used to determine whether a point is in front of a camera or not.

Result 21.6. *The point $\mathbf{X}$ lies in front of the camera $\mathtt{P}$ if and only if* $\mathrm{depth}(\mathbf{X}; \mathtt{P}) > 0$.

In fact, depth is positive for points in front of the camera, negative for points behind the camera, infinite on the plane at infinity and zero on the principal plane of the camera. If the camera centre or the point $\mathbf{X}$ is at infinity, then depth is not defined.

Usually, in this section we will only be concerned with the sign of depth and not its magnitude. We may then write

$$\text{depth}(\mathbf{X}; \mathtt{P}) \doteq w\mathtt{T} \det \mathtt{M} \tag{21.2}$$

where the symbol $\doteq$ indicates equality of sign. The quantity sign(depth($\mathbf{X}; \mathtt{P}$)) will be referred to as the *cheirality* of the point $\mathbf{X}$ with respect to the camera $\mathtt{P}$. The cheirality of a point is said to be reversed by a transformation if it is swapped from 1 to -1 or vice versa.

21.3 Three-dimensional point sets

In this section the connection between cheirality of points with respect to a camera and convex hulls of point sets will be explained. The main result is stated now.

Theorem 21.7. *Let $\mathtt{P}^{\mathrm{E}}$ and $\mathtt{P}'^{\mathrm{E}}$ be two cameras, $\mathbf{X}_i^{\mathrm{E}}$ a set of points lying in front of both cameras, and $\mathbf{x}_i$ and $\mathbf{x}'_i$ the corresponding image points. (The superscript E stands for Euclidean.)*

(i) *Let $(\mathtt{P}, \mathtt{P}', \{\mathbf{X}_i\})$ be any projective reconstruction from the image correspondences $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$, and let $\mathtt{P}\mathbf{X}_i = w_i\hat{\mathbf{x}}_i$ and $\mathtt{P}'\mathbf{X}_i = w'_i\hat{\mathbf{x}}'_i$. Then $w_i w'_i$ has the same sign for all i.*

(ii) *If each $\mathbf{X}_i$ is a finite point, and $\mathtt{P}\widehat{\mathbf{X}}_i = w_i\hat{\mathbf{x}}_i$ with w_i having the same sign for all i, then there exists a quasi-affine transformation $\mathtt{H}$ taking each $\mathbf{X}_i$ to $\mathbf{X}_i^{\mathrm{E}}$.*

Of course, the existence of a projective transformation taking each $\mathbf{X}_i$ to $\mathbf{X}_i^{\mathrm{E}}$ is guaranteed by theorem 10.1(*p*266). The current theorem gives the extra information that the transformation is quasi-affine, and hence one has a quasi-affine reconstruction.

Note that the condition that each $w_i w'_i$ have the same sign is unaffected by multiplying $\mathtt{P}$, $\mathtt{P}'$ or any of the points $\mathbf{X}_i$ by a scale factor, and hence is invariant of the choice of homogeneous representative for any of these quantities. In particular, if $\mathtt{P}$ is multiplied by a negative constant, then so is w_i for all i. Thus the sign of $w_i w'_i$ is inverted for each i, but they still all have the same sign. Similarly, if one point $\mathbf{X}_i$ is multiplied by a negative constant, then both w_i and w'_i change signs, and so the sign of $w_i w'_i$ is unchanged. In the same way, the condition that each w_i (in part (ii) of the theorem) have the same sign for all i is unaffected if the camera matrix is multiplied by a negative (or of course positive) constant.

Proof. The points $\mathbf{X}_i^{\mathrm{E}}$ lie in front of the cameras $\mathtt{P}^{\mathrm{E}}$ and $\mathtt{P}'^{\mathrm{E}}$, and hence have positive depth with respect to these cameras. According to (21.2)

$$\text{depth}(\mathbf{X}_i^{\mathrm{E}}; \mathtt{P}^{\mathrm{E}}) \doteq \det(\mathtt{M}^{\mathrm{E}}) w_i \mathtt{T}_i^{\mathrm{E}}.$$

Hence, $\det(\mathtt{M}^{\mathrm{E}}) w_i \mathtt{T}_i^{\mathrm{E}} > 0$ for all i. Similarly for the second camera, $\det(\mathtt{M}'^{\mathrm{E}}) w'_i \mathtt{T}_i^{\mathrm{E}} > 0$. Multiplying these expressions together, and cancelling $\mathtt{T}_i^{\mathrm{E}}$ because it appears twice, gives $w_i w_i' \det \mathtt{M}^{\mathrm{E}} \det \mathtt{M}'^{\mathrm{E}} > 0$. Since $\det \mathtt{M}^{\mathrm{E}} \det \mathtt{M}'^{\mathrm{E}}$ is constant, this shows that $w_i w_i'$ has constant sign.

This was shown in terms of the true configuration. Note however that for any $\mathtt{H}$, one has $w_i \hat{\mathbf{x}}_i = \mathtt{P}^{\mathrm{E}} \mathbf{X}_i^{\mathrm{E}} = (\mathtt{P}^{\mathrm{E}} \mathtt{H}^{-1})(\mathtt{H}\mathbf{X}_i^{\mathrm{E}})$, and hence $w_i w_i'$ has the same sign for the projective reconstruction $(\mathtt{P}^{\mathrm{E}}\mathtt{H}^{-1}, \mathtt{P}'^{\mathrm{E}}\mathtt{H}^{-1}, \{\mathtt{H}\mathbf{X}_i^{\mathrm{E}}\})$. Since any projective reconstruction is of this form (except for homogeneous scale factors), and the condition that $w_i w_i'$ has the same sign is independent of choice of homogeneous representatives of $\mathbf{X}_i^{\mathrm{E}}$, $\mathtt{P}^{\mathrm{E}}$ and $\mathtt{P}'^{\mathrm{E}}$, it follows that in any projective reconstruction, $w_i w_i'$ has the same sign for all i. This proves the first part of the theorem.

To show the second part, suppose in the projective reconstruction $w_i \hat{\mathbf{x}}_i = \mathtt{P}\widehat{\mathbf{X}}$ all the w_i have the same sign. Since this is a projective reconstruction, there exists a transformation represented by $\mathtt{H}$ such that $\mathtt{H}\widehat{\mathbf{X}}_i = \eta_i \widehat{\mathbf{X}}_i^{\mathrm{E}}$ and $\mathtt{P}\mathtt{H}^{-1} = \epsilon \mathtt{P}^{\mathrm{E}}$ for some constants η_i and ϵ. Then,

$$w_i \hat{\mathbf{x}}_i = \mathtt{P}\widehat{\mathbf{X}}_i = (\mathtt{P}\mathtt{H}^{-1})(\mathtt{H}\widehat{\mathbf{X}}_i) = (\epsilon \mathtt{P}^{\mathrm{E}})(\eta_i \widehat{\mathbf{X}}_i^{\mathrm{E}}) \ .$$

and so

$$\mathtt{P}^{\mathrm{E}} \widehat{\mathbf{X}}_i^{\mathrm{E}} = (w_i/\epsilon\eta_i)\hat{\mathbf{x}}_i$$

for all i. However, since $\operatorname{depth}(\mathbf{X}_i^{\mathrm{E}}, \mathtt{P}^{\mathrm{E}}) > 0$, one has $\det(\mathtt{M}^{\mathrm{E}}) w_i/\epsilon\eta_i > 0$ for all i. Since $\det(\mathtt{M}^{\mathrm{E}})/\epsilon$ is constant, and by hypothesis w_i has the same sign for all i, it follows that η_i has the same sign for all i. Thus the mapping $\mathtt{H}$ such that $\mathtt{H}\widehat{\mathbf{X}}_i = \eta_i \widehat{\mathbf{X}}_i^{\mathrm{E}}$ is a quasi-affine map with respect to the points $\widehat{\mathbf{X}}_i$, according to theorem 21.2.

Note that the condition that w_i have the same sign for all i needs to be checked for one of the cameras only. However, defining $\mathtt{P}'\widehat{\mathbf{X}}_i = w_i' \hat{\mathbf{x}}_i'$, according to part (i) of the theorem, $w_i w_i'$ has the same sign for all i. Thus, if all w_i have the same sign, then so do all w_i'. □

21.4 Obtaining a quasi-affine reconstruction

According to theorem 21.7, any projective reconstruction in which $\mathtt{P}\widehat{\mathbf{X}}_i = w_i \hat{\mathbf{x}}_i$ and w_i has the same sign for all i is a quasi-affine reconstruction. The advantage of a quasi-affine reconstruction is that it gives a closer approximation to the true shape of the object than does an arbitrary projective transformation. It may be used as a stepping stone on the way to a metric reconstruction of the scene, as in [Hartley-94b]. In addition, one may retrieve the convex hull of the object or determine such questions as whether two points lie on the same side of a plane.

It turns out that quasi-affine reconstruction is extremely simple, given a projective reconstruction, as shown in the following theorem.

Theorem 21.8. *Any projective reconstruction in which one of the cameras is an affine camera is a quasi-affine reconstruction.*

Proof. Recall that an affine camera is one for which the last row is of the form $(0, 0, 0, 1)$. In this case, writing $w_i \hat{\mathbf{x}}_i = \mathtt{P}\widehat{\mathbf{X}}_i$, one immediately verifies that $w_i = 1$ for all i, and in particular they all have the same sign. According to theorem 21.2 this means that the reconstruction differs by a quasi-affine transformation from the truth. □

The following result follows immediately.

Corollary 21.9. *Let* $(\mathtt{P}, \mathtt{P}', \{\mathbf{X}_i\})$ *be a projective reconstruction of a scene in which* $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$. *Then by swapping the last two columns of* $\mathtt{P}$ *and of* $\mathtt{P}'$, *as well as the last two coordinates of each* $\mathbf{X}_i$, *one obtains a quasi-affine reconstruction of the scene.*

This is similar to result 10.4(*p*271), in which it was shown that if the camera $\mathtt{P}$ is known in reality to be an affine camera, then the above procedure provides an affine reconstruction.

21.5 Effect of transformations on cheirality

At this point, it is desirable to derive a slightly different form of the formula for depth defined in (21.2). Let $\mathtt{P}$ be a camera matrix. The centre of $\mathtt{P}$ is the unique point $\mathbf{C}$ such that $\mathtt{P}\mathbf{C} = \mathbf{0}$. One can write an explicit formula for $\mathbf{C}$ as follows.

Definition 21.10. Given a camera matrix $\mathtt{P}$, we define $\mathbf{C}_{\mathtt{P}}^{\mathsf{T}}$ to be the vector (c_1, c_2, c_3, c_4), where

$$c_i = (-1)^i \det \hat{\mathtt{P}}^{(i)}$$

and $\hat{\mathtt{P}}^{(i)}$ is the matrix obtained by removing the i-th column of $\mathtt{P}$.

We denote by $[\mathtt{P}/\mathbf{V}^{\mathsf{T}}]$ the 4×4 matrix made up of a 3×4 camera matrix $\mathtt{P}$ augmented with an final row $\mathbf{V}^{\mathsf{T}}$. Definition 21.10 leads to a simple formula for $\det[\mathtt{P}/\mathbf{V}^{\mathsf{T}}]$. Cofactor expansion of the determinant along the last row gives $\det[\mathtt{P}/\mathbf{V}^{\mathsf{T}}] = \mathbf{V}^{\mathsf{T}}\mathbf{C}_{\mathtt{P}}$ for any row vector $\mathbf{V}^{\mathsf{T}}$. As a special case, if $\mathbf{p}_i^{\mathsf{T}}$ is the i-th row of $\mathtt{P}$, then

$$\mathbf{p}_i^{\mathsf{T}}\mathbf{C}_{\mathtt{P}} = \det[\mathtt{P}/\mathbf{p}_i^{\mathsf{T}}] = 0$$

where the last equality is true because the matrix has a repeated row. Since this is true for all i, it follows that $\mathtt{P}\mathbf{C}_{\mathtt{P}} = \mathbf{0}$, and so $\mathbf{C}_{\mathtt{P}}$ is the camera centre, as claimed.

Note that submatrix $\hat{\mathtt{P}}^{(4)}$ is the same as matrix $\mathtt{M}$ in the decomposition $\mathtt{P} = [\mathtt{M} \mid \mathbf{v}]$, and so $\det \mathtt{M} = c_4$. This allows us to reformulate (21.2), as follows.

$$\operatorname{depth}(\mathbf{X}; \mathtt{P}) \doteq w(\mathbf{E}_4^{\mathsf{T}}\mathbf{X})(\mathbf{E}_4^{\mathsf{T}}\mathbf{C}_{\mathtt{P}}) \tag{21.3}$$

where $\mathbf{E}_4^{\mathsf{T}}$ is the vector $(0, 0, 0, 1)$. It is significant to note here that $\mathbf{E}_4$ is the vector representing the plane at infinity – a point $\mathbf{X}$ lies on the plane at infinity if and only if $\mathbf{E}_4^{\mathsf{T}}\mathbf{X} = 0$.

We now consider a projective transformation represented by matrix $\mathtt{H}$. If $\mathtt{P}' = \mathtt{P}\mathtt{H}^{-1}$ and $\mathbf{X}' = \mathtt{H}\mathbf{X}$ then the image correspondences are preserved by this transformation. When speaking of a projective transformation being applied to a set of points and to a camera, it is meant that a point $\mathbf{X}$ is transformed to $\mathtt{H}\mathbf{X}$ and the camera matrix is transformed to $\mathtt{P}\mathtt{H}^{-1}$.

In this section we will consider such projective transformations and their effect on the cheirality of points with respect to a camera. First, we wish to determine what happens to $\mathbf{C}_\mathtt{P}$ when $\mathtt{P}$ is transformed to $\mathtt{PH}^{-1}$. To answer that question, consider an arbitrary 4-vector $\mathbf{V}$. We see that

$$\mathbf{V}^\mathsf{T}\mathtt{H}^{-1}\mathbf{C}_{\mathtt{PH}^{-1}} = \det(\mathtt{PH}^{-1}/\mathbf{V}^\mathsf{T}\mathtt{H}^{-1}) = \det(\mathtt{P}/\mathbf{V}^\mathsf{T})\det\mathtt{H}^{-1} = \mathbf{V}^\mathsf{T}\mathbf{C}_\mathtt{P}\det\mathtt{H}^{-1}.$$

Since this is true for all vectors $\mathbf{V}$, it follows that $\mathtt{H}^{-1}\mathbf{C}_{\mathtt{PH}^{-1}} = \mathbf{C}_\mathtt{P}\det\mathtt{H}^{-1}$, or

$$\mathbf{C}_{\mathtt{PH}^{-1}} = \mathtt{H}\mathbf{C}_\mathtt{P}\det\mathtt{H}^{-1}. \tag{21.4}$$

At one level, this formula is saying that the transformation $\mathtt{H}$ takes the camera centre $\mathbf{C} = \mathbf{C}_\mathtt{P}$ to the new location $\mathbf{C}_{\mathtt{PH}^{-1}} \approx \mathtt{H}\mathbf{C}$. However, we are interested in the exact coordinates of $\mathbf{C}_{\mathtt{PH}^{-1}}$, especially the sign of the last coordinate c_4 which appears in (21.3). Thus, the factor $\det\mathtt{H}^{-1}$ is significant.

Now, applying (21.4) to (21.3) gives

$$\begin{aligned}\text{depth}(\mathtt{H}\mathbf{X};\mathtt{PH}^{-1}) &\doteq w(\mathbf{E}_4^\mathsf{T}\mathtt{H}\mathbf{X})(\mathbf{E}_4^\mathsf{T}\mathbf{C}_{\mathtt{PH}^{-1}})\\ &\doteq w(\mathbf{E}_4^\mathsf{T}\mathtt{H}\mathbf{X})(\mathbf{E}_4^\mathsf{T}\mathtt{H}\mathbf{C}_\mathtt{P})\det\mathtt{H}^{-1} \quad .\end{aligned}$$

One may interpret the expression $\mathbf{E}_4^\mathsf{T}\mathtt{H}$ as being the plane $\boldsymbol{\pi}_\infty$ mapped to infinity by $\mathtt{H}$. This is because a point $\mathbf{X}$ lies on $\boldsymbol{\pi}_\infty$ if and only if the last coordinate of $\mathtt{H}\mathbf{X}$ is zero – that is $\mathbf{E}_4\mathtt{H}\mathbf{X} = 0$. On the other hand, $\mathbf{X}$ lies on $\boldsymbol{\pi}_\infty$ if and only if $\boldsymbol{\pi}_\infty^\mathsf{T}\mathbf{X} = 0$. Finally, denoting the fourth row of the transformation matrix $\mathtt{H}$ by $\mathbf{h}_4^\mathsf{T}$, and $\text{sign}(\det\mathtt{H})$ by δ, we obtain

Result 21.11. *If $\boldsymbol{\pi}_\infty$ is the plane mapped to infinity by a projective transformation $\mathtt{H}$ and $\delta = \text{sign}(\det\mathtt{H})$, then*

$$\text{depth}(\mathtt{H}\mathbf{X};\mathtt{PH}^{-1}) \doteq w(\boldsymbol{\pi}_\infty^\mathsf{T}\mathbf{X})(\boldsymbol{\pi}_\infty^\mathsf{T}\mathbf{C}_\mathtt{P})\delta \quad .$$

This equation will be used extensively. It may be considered to be a generalization of (21.3). It will be seen in the next section that $\delta = \text{sign}(\det\mathtt{H})$ is an indicator of whether $\mathtt{H}$ is an orientation-reversing or orientation-preserving transformation. Thus, the effect on cheirality of applying a transformation $\mathtt{H}$ is determined only by the position of the plane mapped to infinity $\boldsymbol{\pi}_\infty$, and whether $\mathtt{H}$ preserves or reverses orientation.

We now consider the effect of different transformations on the cheirality of points with respect to a camera. The effect of an affine transformation is considered first.

Result 21.12. *An affine transformation with positive determinant preserves the cheirality of any point with respect to a camera. An affine transformation with negative determinant reverses cheirality.*

Proof. An affine transformation preserves the plane at infinity, hence $\boldsymbol{\pi}_\infty = \mathbf{E}_4$. The result then follows by comparing (21.3) and result 21.11. □

We now determine how an arbitrary projective transformation affects cheirality.

Result 21.13. *Let $\mathtt{H}$ represent a projective transformation with positive determinant, and let $\boldsymbol{\pi}_\infty$ be the plane in space mapped to infinity by $\mathtt{H}$. The cheirality of a point $\mathbf{X}$ is preserved by $\mathtt{H}$ if and only if $\mathbf{X}$ lies on the same side of the plane $\boldsymbol{\pi}_\infty$ as the camera centre.*

Proof. Since $\det \mathtt{H} > 0$, we see from (21.3) and result 21.11 that $\text{depth}(\mathbf{X}; \mathtt{P}) \doteq \text{depth}(\mathtt{H}\mathbf{X}; \mathtt{P}\mathtt{H}^{-1})$ if and only if $(\boldsymbol{\pi}_\infty^\mathsf{T}\mathbf{X})(\boldsymbol{\pi}_\infty^\mathsf{T}\mathbf{C}) \doteq (\mathbf{E}_4^\mathsf{T}\mathbf{X})(\mathbf{E}_4^\mathsf{T}\mathbf{C})$. Suppose the point $\mathbf{X}$ and the camera $\mathtt{P}$ are located at finite points so that the cheirality is well defined, and let them be scaled so that both $\mathbf{X}$ and $\mathbf{C}$ have final coordinate equal to 1. In this case, $(\mathbf{E}_4^\mathsf{T}\mathbf{X})(\mathbf{E}_4^\mathsf{T}\mathbf{C}) = 1$ and we see that cheirality is preserved if and only if $(\boldsymbol{\pi}_\infty^\mathsf{T}\mathbf{X})(\boldsymbol{\pi}_\infty^\mathsf{T}\mathbf{C}) \doteq 1$, or otherwise expressed $\boldsymbol{\pi}_\infty^\mathsf{T}\mathbf{X} \doteq \boldsymbol{\pi}_\infty^\mathsf{T}\mathbf{C}$. This condition may be interpreted as meaning that the points $\mathbf{C}$ and $\mathbf{X}$ both lie on the same side of the plane $\boldsymbol{\pi}_\infty$. Hence, the cheirality of a point $\mathbf{X}$ is preserved by a transformation $\mathtt{H}$, if and only if it lies on the same side of the plane $\boldsymbol{\pi}_\infty$ as the camera centre. □

21.6 Orientation

We now consider the question of image orientation. A mapping h from $\mathbb{R}^n$ to itself is called orientation-preserving at points $\mathbf{X}$ where the Jacobian of h (the determinant of the matrix of partial derivatives) is positive, and orientation-reversing at points where the Jacobian is negative. Reflection of points of $\mathbb{R}^n$ with respect to a hyperplane (that is mirror imaging) is an example of an orientation-reversing mapping. A projectivity h from $\mathtt{P}^n$ to itself restricts to a mapping from $\mathbb{R}^n - \boldsymbol{\pi}_\infty$ to $\mathbb{R}^n$, where $\boldsymbol{\pi}_\infty$ is the hyperplane (line, plane) mapped to infinity by $\mathtt{H}$. Consider the case $n = 3$ and let $\mathtt{H}$ be a 4×4 matrix representing the projectivity h. We wish to determine at which points $\mathbf{X}$ in $\mathbb{R}^n - \boldsymbol{\pi}_\infty$ the map h is orientation-preserving. It may be verified (quite easily using Mathematica [Mathematica-92]) that if $\mathtt{H}\widehat{\mathbf{X}} = w\widehat{\mathbf{X}}'$ and $\mathtt{J}$ is the Jacobian matrix of h evaluated at $\mathbf{X}$, then $\det(\mathtt{J}) = \det(\mathtt{H})/w^4$. This gives the following result.

Result 21.14. *A projectivity h of $\mathbb{P}^3$ represented by a matrix $\mathtt{H}$ is orientation-preserving at any point in $\mathbb{R}^3 - \boldsymbol{\pi}_\infty$ if and only if $\det(\mathtt{H}) > 0$.*

Of course, the concept of orientability may be extended to the whole of $\mathbb{P}^3$, and it may be shown that h is orientation-preserving on the whole of $\mathbb{P}^3$ if and only if $\det(\mathtt{H}) > 0$. The essential feature here is that as a topological manifold, $\mathbb{P}^3$ is orientable.

Two sets of points $\{\mathbf{X}_i\}$ and $\{\overline{\mathbf{X}}_i\}$ that correspond via a quasi-affine transformation are said to be *oppositely oriented* if the transformation is orientation-reversing. As an example, consider the transformation given by a diagonal matrix $\mathtt{H} = \text{diag}(1, 1, -1, 1)$. This transformation has negative determinant, and hence is orientation-reversing. On the other hand, it is affine, and hence quasi-affine. Therefore, it is possible always to construct oppositely oriented quasi-affine reconstructions of a scene. It may appear therefore that the orientation of a scene may not be determined from a pair of images. Although this is sometimes true, it is sometimes possible to rule out one of the oppositely oriented quasi-affine reconstructions of a scene, and hence determine the true orientation of the scene.

Common experience provides some clues here. In particular a stereo pair may be viewed by presenting one image to one eye and the other image to the other eye. If this is done correctly, then the brain perceives a 3D reconstruction of the scene If, however, the two images are swapped and presented to the opposite eyes, then the perspective will be reversed – hills become valleys and vice versa. In effect, the brain is able to compute two oppositely oriented reconstructions of the image pair. It seems, therefore,

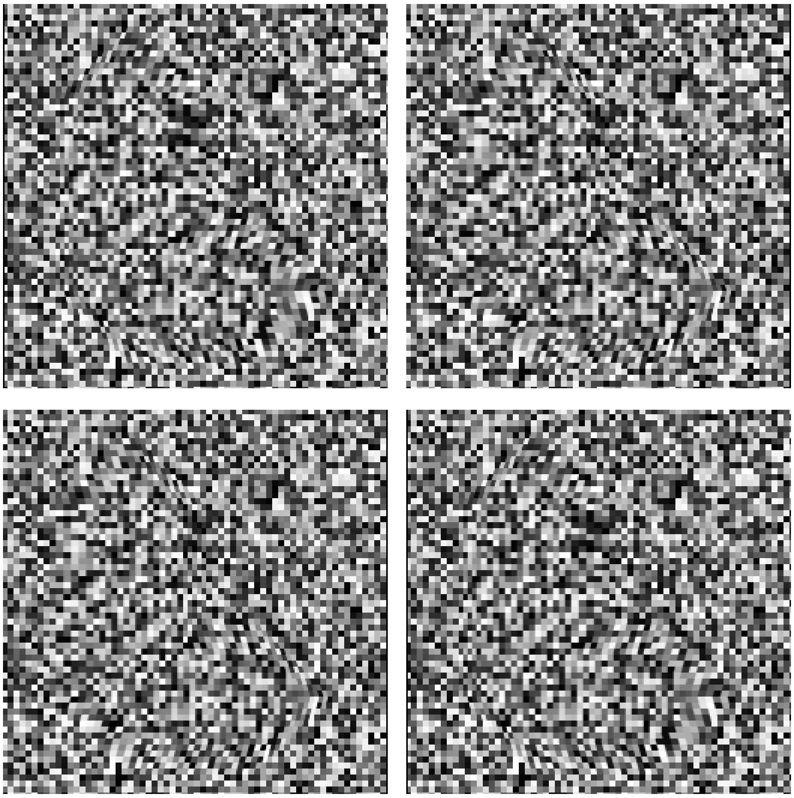


Fig. 21.2. *Stereo pairs of images that may be viewed by cross-fusing (the eyes are crossed so that the left eye sees the right image and vice versa). The two bottom images are the same as the top pair, except that they are swapped. In the top pair of images one sees an L-shaped region raised above a planar background. In the bottom pair of images the L-shaped region appears as an indentation. The two "reconstructions" differ by reflection in the background plane. This demonstrates that the same pair of images may give rise to two differently oriented projective reconstructions.*

that in certain circumstances, two oppositely oriented realizations of an image pair exist. This is illustrated in figure 21.2.

It may be surprising to discover that this is not always the case, as is shown in the following theorem. As used in this theorem and elsewhere in this chapter, a projective realization of a set of point correspondences is known as a *strong realization* if the reconstructed 3D points $\mathbf{X}_i$ are in front of all the cameras.

Theorem 21.15. *Let* $(\mathtt{P}, \mathtt{P}', \{\mathbf{X}_i\})$ *be a strong realization of a uniquely realizable set of point correspondences. There exists a different oppositely oriented strong realization* $(\overline{\mathtt{P}}, \overline{\mathtt{P}}', \{\overline{\mathbf{X}}_i\})$ *if and only if there exists a plane in* $\mathbb{R}^3$ *such that the perspective centres of both cameras* $\mathtt{P}$ *and* $\mathtt{P}'$ *lie on one side of the plane, and the points* $\mathbf{X}_i$ *lie on the other side.*

Proof. Consider one strong realization of the configuration. By definition, all the points lie in front of both cameras. Suppose that there exists a plane separating the points from the two camera centres. Let $\mathtt{G}$ be a projective transformation mapping the given plane to infinity and let $\mathtt{A}$ be an affine transformation. Suppose further that

$\det \mathtt{G} > 0$ and $\det \mathtt{A} < 0$. Let $\mathtt{H}$ be the composition $\mathtt{H} = \mathtt{AG}$. According to result 21.13 the transformation $\mathtt{G}$ is cheirality-reversing for the points, since the points are on the opposite side of the plane from the camera centres. According to result 21.12 $\mathtt{A}$ is also cheirality-reversing, since $\det \mathtt{A} < 0$. The composition $\mathtt{H}$ must therefore be cheirality-preserving, and it transforms the strong configuration to a different strong configuration. Since $\mathtt{H}$ has negative determinant, however, it is orientation-reversing, so the two strong realizations have opposite orientations.

Conversely, suppose that two strong oppositely oriented realizations exist and let $\mathtt{H}$ be the transformation taking one to the other. Since $\mathtt{H}$ is orientation-reversing, $\det \mathtt{H} < 0$. The mapping $\mathtt{H}$ is by definition cheirality-preserving on all points, with respect to both cameras. If $\boldsymbol{\pi}_\infty$ is the plane mapped to infinity by $\mathtt{H}$, then according to result 21.13 the points $\mathbf{X}$ must lie on the opposite side of the plane $\boldsymbol{\pi}_\infty$ from both camera centres. □

21.7 The cheiral inequalities

In section 21.4 a very simple method was given for obtaining a quasi-affine reconstruction of a scene directly from a projective reconstruction. However, the reconstruction obtained there did not respect the condition that the points must lie in front of all cameras. In fact, the first camera in this construction is an affine camera, for which front and back are not well defined. By taking full advantage of the fact that visible points must lie in front of the camera, it is possible to constrain the reconstruction more tightly, leading to a closer approximation to a true affine reconstruction of the scene.

The method will be given for the case of a reconstruction derived from several images. One is given a set of image points $\{\mathbf{x}_i^j\}$, where $\mathbf{x}_i^j$ is the projection of the i-th point in the j-th image. Not all points are visible in each image, so for some (i, j) the point $\mathbf{x}_i^j$ is not given, in which case it is not known whether the i-th point lies in front of the j-th camera or not. On the other hand, the existence of an image point $\mathbf{x}_i^j$ implies that the point lies in front of the camera.

We start from an assumed projective reconstruction of the scene, consisting of a set of 3D points $\mathbf{X}_i$ and cameras $\mathtt{P}^j$ such that $\mathbf{x}_i^j \approx \mathtt{P}^j\mathbf{X}_i$. Writing the implied scalar constant explicitly in this equation gives $w_i^j\hat{\mathbf{x}}_i^j = \mathtt{P}^j\mathbf{X}_i$. In this equation, $\mathtt{P}^j$ and $\mathbf{X}_i$ are arbitrarily chosen homogeneous representatives of the respective matrix or vector. Related to theorem 21.7 one may state for several views:

Result 21.16. *Consider a set of points $\mathbf{X}_i^{\mathrm{E}}$ and cameras $\mathtt{P}^{j\mathrm{E}}$, and let $\mathbf{x}_i^j = \mathtt{P}^{j\mathrm{E}}\mathbf{X}_i^{\mathrm{E}}$ be defined for some indices (i, j) such that point $\mathbf{X}_i^{\mathrm{E}}$ lies in front of camera $\mathtt{P}^{j\mathrm{E}}$. Let $(\mathtt{P}^j; \mathbf{X}_i)$ be a projective reconstruction from $\mathbf{x}_i^j$. Then there are camera matrices $\tilde{\mathtt{P}}^j = \pm\mathtt{P}^j$ and $\tilde{\mathbf{X}}_i = \pm\mathbf{X}_i$ such that for each (i, j) for which $\mathbf{x}_i^j$ is defined, one has*

$$\tilde{\mathtt{P}}^j\tilde{\mathbf{X}}_i = w_i^j\hat{\mathbf{x}}_i^j \;\; \textit{with} \;\; w_i^j > 0.$$

Briefly stated, one can always adjust a projective reconstruction, multiplying camera matrices and points by -1 if necessary, so that w_i^j is positive whenever image point $\mathbf{x}_i^j$ exists. The simple proof is omitted. To find the matrices $\tilde{\mathtt{P}}^j$ and points $\tilde{\mathbf{X}}_i$, one may assume that one of the cameras $\tilde{\mathtt{P}}_1 = \mathtt{P}_1$, for otherwise all points and cameras may be multiplied by -1. The condition $\mathtt{P}_1\tilde{\mathbf{X}}_i = w_i^1\mathbf{x}_i^1$ with $w_i^1 > 0$ determines whether to choose $\tilde{\mathbf{X}}_i = \mathbf{X}_i$ or $-\mathbf{X}_i$ for all i such that $\mathbf{x}_i^1$ is defined. Each known $\tilde{\mathbf{X}}_i$ determines $\tilde{\mathtt{P}}^j$ for all j such that $\mathbf{x}_i^j$ is defined. Continuing in this way, one easily finds the factor ± 1 to apply to each $\mathtt{P}^j$ and $\mathbf{X}_j$ to find $\tilde{\mathtt{P}}^j$ and $\tilde{\mathbf{X}}_i$. We assume that this has been done, and replace each $\mathtt{P}^j$ by $\tilde{\mathtt{P}}^j$ and $\mathbf{X}_i$ by $\tilde{\mathbf{X}}_i$. In future we drop the tildes and continue to work with the corrected $\mathtt{P}^j$ and $\mathbf{X}_i$. We now know that $w_i^j > 0$ whenever image point $\mathbf{x}_i^j$ is given.

Now, we seek a transformation $\mathtt{H}$ that will transform the projective reconstruction to a quasi-affine reconstruction for which all points lie in front of the cameras as appropriate. Denoting by 4-vector $\mathbf{v}$ the plane $\boldsymbol{\pi}_\infty$ mapped to infinity by $\mathtt{H}$, this condition may be written as (see result 21.11):

$$\text{depth}(\mathbf{X}_i; \mathtt{P}^j) \doteq (\mathbf{v}^\mathsf{T}\mathbf{X}_i)(\mathbf{v}^\mathsf{T}\mathbf{C})\delta > 0$$

where $\delta = \text{sign}(\det \mathtt{H})$. This condition holds for all pairs (i, j) where $\mathbf{x}_i^j$ is given.

Since we are free to multiply $\mathbf{v}$ by -1 if necessary, we may assume that $(\mathbf{v}^\mathsf{T}\mathbf{C}^1)\delta > 0$ for the centre of the camera $\mathtt{P}^1$. The following inequalities now follow easily:

$$\begin{aligned} \mathbf{X}_i^\mathsf{T}\mathbf{v} > 0 \quad &\text{for all } i \\ \delta \mathbf{C}^{j\mathsf{T}}\mathbf{v} > 0 \quad &\text{for all } j. \end{aligned} \tag{21.5}$$

These equations (21.5) may be called the *cheiral inequalities*. Since the values of each $\mathbf{X}_i$, $\mathbf{C}$ and $\mathbf{C}'$ are known, they form a set of inequalities in the entries of $\mathbf{v}$. The value of δ is not known *a priori*, and so it is necessary to seek a solution for each of the two cases $\delta = 1$ and $\delta = -1$.

To find the required transformation $\mathtt{H}$, first of all we solve the cheiral inequalities to find a value of $\mathbf{v}$, either for $\delta = 1$ or $\delta = -1$. The required matrix $\mathtt{H}$ is any matrix having $\mathbf{v}^\mathsf{T}$ as its last row and satisfying the condition $\det \mathtt{H} \doteq \delta$. If the last component of $\mathbf{v}$ is non-zero, then $\mathtt{H}$ can be chosen to have the simple form in which the first three rows are of the form $\pm[\mathtt{I} \mid \mathbf{0}]$.

If a Euclidean reconstruction (or more specifically a quasi-affine reconstruction) is possible, then there must be a solution either for $\delta = 1$ or $\delta = -1$. In some cases there will exist solutions of the cheiral inequalities for both $\delta = 1$ and $\delta = -1$. This will mean that two oppositely oriented strong realizations exist. The conditions under which this may occur were discussed in section 21.6.

Solving the cheiral inequalities

Naturally, the cheiral inequalities may be solved using techniques of linear programming. As they stand however, if $\mathbf{v}$ is a solution then so is $\alpha\mathbf{v}$ for any positive factor α. In order to restrict the solution domain to be bounded, one may add additional inequalities. For instance, if $\mathbf{v} = (v_1, v_2, v_3, v_4)^\mathsf{T}$, then the inequalities $-1 < v_i < 1$ serve to restrict the solution domain to be a bounded polyhedron.

To achieve a unique solution we need to specify some goal function to be linearized. An appropriate strategy is to seek to maximize the extent to which each of the inequalities is satisfied. To do this, we introduce one further variable, d. Each of the inequalities $\mathbf{a}^\mathsf{T}\mathbf{v}$ of the form (21.5) for appropriate $\mathbf{a}$ is replaced by an inequality $\mathbf{a}^\mathsf{T}\mathbf{v} > d$. We seek to maximize d while satisfying all the inequalities. This is a standard linear programming problem, for which many methods of solution exist, such as the simplex method ([Press-88]).[1] If a solution is found for which $d > 0$ then this will be a desired solution.

Summary of the algorithm

Now, we give the complete algorithm for computing a quasi-affine reconstruction of a scene using the cheiral inequalities. The algorithm as outlined above was discussed for the case of two views. In the present case it will be presented for an arbitrary number of views. The extension to more views is straightforward.

Objective

Given a set of 3D points $\mathbf{X}_i$ and camera matrices $\mathtt{P}^j$ constituting a projective reconstruction from a set of image points, compute a projective transformation $\mathtt{H}$ transforming the projective to a quasi-affine reconstruction.

Algorithm

(i) For each pair (i, j) such that point $\mathbf{x}_i^j$ is given let $\mathtt{P}^j \mathbf{X}_i = w_i^j \hat{\mathbf{x}}_i^j$.
(ii) Replace some cameras $\mathtt{P}^j$ by $-\mathtt{P}^j$ and some points $\mathbf{X}_i$ by $-\mathbf{X}_i$ as required to ensure that each $w_i^j > 0$.
(iii) Form the cheiral inequalities (21.5) where $\mathbf{C}^j = \mathbf{C}_{\mathtt{P}^j}$ is defined by definition 21.10.
(iv) For each of the values $\delta = \pm 1$, choose a solution (if it exists) to the set of cheiral inequalities. Let the solution be $\mathbf{v}_\delta$. A solution must exist for at least one value of δ, sometimes for both values of δ.
(v) Define a matrix $\mathtt{H}_\delta$ having last row equal to $\mathbf{v}_\delta$ and such that $\det(\mathtt{H}) \doteq \delta$. The matrix $\mathtt{H}_\delta$ is the required transformation matrix. If both $\mathtt{H}_+$ and $\mathtt{H}_-$ exist, then they lead to two oppositely oriented quasi-affine reconstructions.

Algorithm 21.1. *Computing a quasi-affine reconstruction.*

Bounding the plane at infinity

A quasi-affine reconstruction is of course not unique, being defined only up to a quasi-affine transformation with respect to the points and camera centres. However, once one has been found, it is possible to set bounds on the coordinates of the plane at infinity. Thus, let $\mathtt{P}^j$ and $\mathbf{X}_i$ constitute a quasi-affine reconstruction of a scene. One may choose the sign of $\mathtt{P}^j$ and $\mathbf{X}_i$ such that the last coordinates of $\mathtt{X}_i$ and the determinants of each $\mathtt{M}^j$ are positive. One may apply a translation to the points and cameras so that the coordinate origin lies inside the convex hull of the points and camera centres. For simplicity, the centroid of these points may be placed at the origin.

It is possible to apply a further quasi-affine transformation $\mathtt{H}$ to obtain an alternative reconstruction. Let $\boldsymbol{\pi}_\infty$ be the plane mapped to infinity by $\mathtt{H}$. We confine our interest to orientation-preserving transforms, and wish to find constraints on the coordinates of $\boldsymbol{\pi}_\infty$ such that $\mathtt{H}$ is quasi-affine. A plane $\boldsymbol{\pi}_\infty$ has this property if and only if it lies entirely outside the convex hull of the points and camera centres. Since the plane $\boldsymbol{\pi}_\infty$

[1] The Simplex algorithm given in [Press-88] is not suitable for use as stands, since it makes the unnecessary assumption that all variables are non-negative. It needs to be modified to be used for this problem.

cannot cross the convex hull, it cannot pass through the origin. Representing $\boldsymbol{\pi}_\infty$ by the vector $\mathbf{v}$, this says that the last coordinate of $\mathbf{v}$ is non-zero. One may then write $\mathbf{v} = (v_1, v_2, v_3, 1)^\mathsf{T}$. Since the origin lies on the same side of the plane as all the points, the cheiral inequalities become

$$\begin{aligned} \mathbf{X}_i^\mathsf{T}\mathbf{v} &> 0 \ \text{ for all } i \\ \mathbf{C}^{j\mathsf{T}}\mathbf{v} &> 0 \ \text{ for all } j \end{aligned} \tag{21.6}$$

for $\mathbf{v} = (v_1, v_2, v_3, 1)^\mathsf{T}$. One may find upper and lower bounds for each v_i by solving the linear programming problem to maximize v_i or $-v_i$ subject to these constraints. None of the v_i can be unbounded, since otherwise the plane $\boldsymbol{\pi}_\infty$ represented by the vector $\mathbf{v}$ could lie arbitrarily close to the origin.

Before solving for this system, good practice is to apply an affine transformation to normalize the set of points and camera centres so that their centroid is at the origin and their principal moments are all equal to 1.

The complete algorithm for computing the bounds on the position of the plane at infinity is given in algorithm 21.2.

Objective

Given a quasi-affine reconstruction of a scene, establish bounds on the coordinates of the plane at infinity.

Algorithm

(i) Normalize the points $\mathbf{X}_i = (\mathrm{X}_i, \mathrm{Y}_i, \mathrm{Z}_i, \mathrm{T}_i)^\mathsf{T}$ so that $\mathrm{T}_i = 1$, and cameras $\mathtt{P}^j = [\mathtt{M}_j \mid \mathbf{t}^j]$ so that $\det \mathtt{M}^j = 1$.

(ii) Further normalize by replacing $\mathbf{X}_i$ by $\mathtt{H}^{-1}\mathbf{X}_i$ and $\mathtt{P}^j$ by $\mathtt{P}^j\mathtt{H}$, where $\mathtt{H}$ is an affine transformation moving the centroid to the origin and scaling in the principal axis directions so that is has equal principal axes.

(iii) Letting $\mathbf{v} = (v_1, v_2, v_3, 1)^\mathsf{T}$, form cheiral inequalities (21.6). Any orientation-preserving transformation $\mathtt{H}$ mapping the reconstruction to an affine reconstruction of the image must have the form

$$\mathtt{H} = \begin{bmatrix} & \mathtt{I} & & \mathbf{0} \\ v_1 & v_2 & v_3 & 1 \end{bmatrix}$$

for a vector $\mathbf{v}$ satisfying these inequalities.

(iv) Upper and lower bounds for each v_i may be found by running a linear programming problem six times. The coordinates of a desired transformation $\mathtt{H}$ must lie inside the box defined by these bounds.

Algorithm 21.2. *Establishing bounds on the plane at infinity.*

21.8 Which points are visible in a third view

Consider a scene reconstructed from two views. We consider now the question of determining which points are visible in a third view. Such a question arises when one is given two uncalibrated views of a scene and one seeks to synthesize a third view. This can be done by carrying out a projective reconstruction of the scene from the

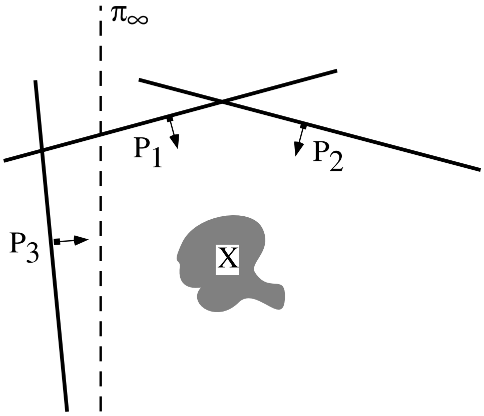


Fig. 21.3. *The point set* $\mathbf{X}$ *is in front of all three cameras as shown. However, if an orientation-preserving projective transformation* $\mathtt{H}$ *is applied taking the plane* $\boldsymbol{\pi}_\infty$ *to infinity, then the point set will subsequently lie in front of the cameras* $\mathtt{P}^1$ *and* $\mathtt{P}^2$ *but behind the camera* $\mathtt{P}^3$*. Thus, suppose the point set* $\mathbf{X}$ *is reconstructed from images captured by cameras* $\mathtt{P}^1$ *and* $\mathtt{P}^2$*, and let* $\mathtt{P}^3$ *be any other camera matrix. If a plane exists separating the centre of camera* $\mathtt{P}^3$ *from the other camera centres, and not meeting the convex hull of the point set* $\mathbf{X}$*, then it cannot be determined whether the points lie in front of* $\mathtt{P}^3$.

first two views and then projecting into the third view. In this case, it is important to determine if a point lies in front of the third camera and is hence visible, or not.

If the third view is given simply by specifying the camera matrix with respect to the frame of reference of some given reconstruction, then it may be impossible to determine whether points are in front of the third camera or behind it in the true scene. The basic ambiguity is illustrated in figure 21.3. Knowledge of a single point known to be visible in the third view serves to break the ambiguity, however, as the following result shows. By applying theorem 21.7(*p*519) to the first and third views, one obtains the following criterion.

Result 21.17. *Let points* $(\mathtt{P}^1, \mathtt{P}^2, \{\mathbf{X}_i\})$ *be a realization of a set of correspondences* $\mathbf{x}_i^1 \leftrightarrow \mathbf{x}_i^2$*. Let* $\mathtt{P}^3$ *be the camera matrix of a third view and suppose that* $w_j^i \hat{\mathbf{x}}_i = \mathtt{P}^i \mathbf{X}_j$ *for* $i = 1, \ldots, 3$*. Then* $w_j^1 w_j^3$ *has the same sign for all points* $\mathbf{X}_j$ *visible in the third view.*

In practice, it will usually be the case that one knows at least one point $\mathbf{X}_0$ visible in the third view. This serves to define the sign $w_0^1 w_0^3$, and any other point $\mathbf{X}_j$ will be in front of the camera $\mathtt{P}^3$ if and only if $w_j^1 w_j^3 \doteq w_0^1 w_0^3$.

As an example, once a projective reconstruction has been carried out using two views, the camera matrix of the third camera may be determined from the images of six or more points known to be in front of it by solving directly for the matrix $\mathtt{P}^3$ given the correspondences $\mathbf{x}_i^3 = \mathtt{P}^3 \mathbf{X}_i$ where points $\mathbf{X}_i$ are the reconstructed points. Then one can determine unambiguously which other points are in front of $\mathtt{P}^3$.

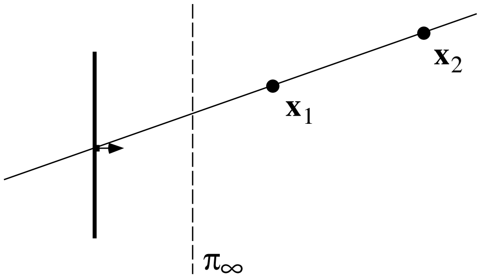


Fig. 21.4. *As shown, the point* $\mathbf{x}_1$ *is closer to the camera than* $\mathbf{x}_2$. *However, if an orientation-reversing projectivity is applied, taking* $\boldsymbol{\pi}_\infty$ *to infinity, then both* $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *remain in front of the camera, but* $\mathbf{x}_2$ *is closer to the camera than* $\mathbf{x}_1$.

21.9 Which points are in front of which

When we are attempting to synthesize a new view of a scene that has been reconstructed from two or more uncalibrated views it is sometimes necessary to consider the possibility of points being obscured by other points. This leads to the question: given two points that project to the same point in the new view, which one is closer to the camera, and hence obscures the other? In the case where the possibility exists of oppositely oriented quasi-affine reconstructions it may once again be impossible to determine which of a pair of points is closer to the new camera. This is illustrated in figure 21.4. If a plane exists, separating the camera centres from the point set, then two oppositely oriented reconstructions exist, and one cannot determine which points are in front of which. The sort of ambiguity shown in figure 21.4 can only occur in the case where there exists a plane $\boldsymbol{\pi}_\infty$ that separates the camera centres from the set of all visible points. If this is not the case, then one can compute a quasi-affine reconstruction and the problem is easily solved. To avoid the effort of computing a quasi-affine reconstruction, however, we would like to solve this problem using only a projective reconstruction of the scene. How this may be done is explained next.

One may invert (21.1–*p*518) to get an expression for $\text{depth}^{-1}(\mathbf{X}; \mathtt{P}) = 1/\text{depth}(\mathbf{X}; \mathtt{P})$. This inverse depth function is infinite on the principal plane of the camera, zero on the plane at infinity, positive for points in front of the camera and negative for points behind the camera. For notational simplicity, we write $\chi(\mathbf{X}; \mathtt{P})$ instead of $\text{depth}^{-1}(\mathbf{X}; \mathtt{P})$.

For points $\mathbf{X}$ lying on a ray through the camera centre, the value of $\chi(\mathbf{X}; \mathtt{P})$ decreases monotonically along this ray, from zero at the camera centre, decreasing through positive values to zero at the plane at infinity, thence continuing to decrease through negative values to $-\infty$ at the camera centre. A point $\mathbf{X}_1$ lies closer to the front of the camera than $\mathbf{X}_2$ if and only if $\chi(\mathbf{X}_1) > \chi(\mathbf{X}_2)$. This is illustrated in figure 21.5.

Now, if the configuration undergoes an orientation-preserving transformation $\mathtt{H}$ taking the plane $\boldsymbol{\pi}_\infty$ to infinity, then the parameter χ will be replaced by a new parameter χ' defined by $\chi'(\mathbf{X}) = \chi(\mathtt{H}\mathbf{X}; \mathtt{PH}^{-1})$. The value of χ' must also vary monotonically along the ray. Since $\mathtt{H}$ is orientation-preserving, points just in front of the camera centre will remain in front of the camera after the transformation, because of result 21.13.

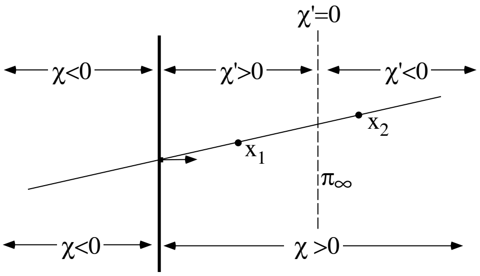


Fig. 21.5. *The parameter χ.*

Thus both χ and χ' decrease monotonically in the same direction along the ray. If $\mathbf{X}_1$ and $\mathbf{X}_2$ are two points on the line, then $\chi'(\mathbf{X}_1) > \chi'(\mathbf{X}_2)$ if and only if $\chi(\mathbf{X}_1) > \chi(\mathbf{X}_2)$.

In the case where the projective transformation has negative determinant, then the front and back of the camera are reversed locally. In this case the direction of increase of the parameter χ' will be reversed. Consequently, $\chi'(\mathbf{X}_1) > \chi'(\mathbf{X}_2)$ if and only if $\chi(\mathbf{X}_1) < \chi(\mathbf{X}_2)$.

In the case where the projective transformation transforms the scene to the "true" scene, of two points that project to the same point in the image, the one with the higher value of χ' is closer to the camera. This leads to the following result that allows us to determine from an arbitrary projective reconstruction which of two points is closer to the front of the camera.

Result 21.18. *Suppose that $\mathbf{X}_1$ and $\mathbf{X}_2$ are two points that map to the same point in an image. Consider a projective reconstruction of the scene and let the parameter χ be defined (by formula (21.3–p521)) in the frame of the projective reconstruction. If the projective reconstruction has the same orientation as the true scene, then the point that lies closer to the front of the camera in the true scene is the one that has the greater value of χ. On the other hand, if the projective transformation has the opposite orientation, then the point with smaller value of χ will lie closer to the front of the camera in the true scene.*

As remarked previously, unless there exists a plane separating the point set from the cameras used for the reconstruction, the orientation of the scene is uniquely determined, and one can determine whether the projective transformation of result 21.18 has positive or negative determinant. However, to do this may require one to compute a strong realization of the configuration by the linear programming method as described in section 21.7. If differently oriented strong realizations exist, then as illustrated by figure 21.4, there is an essential ambiguity. However, this ambiguity may be resolved by knowledge of the relative distance from the camera of a single pair of points.

21.10 Closure

21.10.1 The literature

The topic of this chapter belongs to *Oriented projective geometry*, which is treated in a standard and readable text [Stolfi-91]. Laveau and Faugeras apply the ideas of oriented projective geometry in [Laveau-96b]. The concepts of front and back of the camera were also used in [Robert-93] to compute convex hulls in projective reconstructions. This chapter derives from the paper [Hartley-98a] which also treats such topics of invariants for quasi-affine mappings and conditions under which arbitrary correspondences allow quasi-affine reconstructions.

Cheirality, and specifically the cheirality inequalities have been useful in determining quasi-affine reconstructions as an intermediate step towards affine and metric reconstruction in [Hartley-94b, Hartley-99]. More recently Werner and Pajdla in [Werner-01] have used oriented projective geometry to eliminate spurious line correspondences, and to constrain correspondences of five points [Werner-03], over two views.