Cheirality

Richard I. Hartley
GE - Corporate Research and Development,
P.O. Box 8, Schenectady, NY, 12301.
Ph : (518)-387-7333
Fax : (518)-387-6845
email : hartley@crd.ge.com

June 21, 2001

Abstract

It is known that a set of points in 3 dimensions is determined up to projectivity from two views with uncalibrated cameras. It is shown in this paper that this result may be improved by distinguishing between points in front of and behind the camera. Any point that lies in an image must lie in front of the camera producing that image. Using this idea, it is shown that the scene is determined from two views up to a more restricted class of mappings known as quasi-affine transformations, *which are precisely those projectivities that preserve the convex hull of an object of interest. An invariant of quasi-affine transformation known as the* cheiral sequence *of a set of points is defined and it is shown how the cheiral sequence may be computed using two uncalibrated views. As demonstrated theoretically and by experiment the cheiral sequence may distinguish between sets of points that are projectively equivalent. These results lead to necessary and sufficient conditions for a set of corresponding pixels in two images to be realizable as the images of a set of points in 3 dimensions.*

Using similar methods, a necessary and sufficient condition is given for the orientation *of a set of points to be determined by two views. If the perspective centres are not separated from the point set by a plane, then the orientation of the set of points is determined from two views.*

1 Introduction

Consider a set of points $\{\mathbf{x}_i\}$ lying in a plane in space and let $\{\mathbf{u}_i\}$ and $\{\mathbf{u}'_i\}$ be two images of these points taken with arbitrary uncalibrated perspective (pinhole) cameras. It is well known that the points $\mathbf{u}_i$ and $\mathbf{u}'_i$ are related by a planar projectivity, that is, there exists h a projectivity of the plane such that $h\mathbf{u}_i = \mathbf{u}'_i$ for all i. This fact has been used for the recognition of planar objects. For instance in [13] planar projective invariants were used to define indexing functions allowing look-up of the objects in an object data-base. Since the indexing functions are invariant under plane projectivities, they provide the same value independent of the view of the object.

In a similar way, it has been shown in [1] and [3] that a set of points in 3-dimensions is determined up to a 3-dimensional projectivity by two images taken with uncalibrated cameras. Both these papers give a constructive method for determining the point configuration (up to projectivity). This permits the computation of projective invariants of sets of points seen in two views. An experimental verification of these results has been reported in [2] and is summarized in this paper.

The papers just cited make no distinction between points that lie in front of the camera and those that lie behind. The property of a point that specifies that it lies in front of or behind a given camera will be termed the *cheirality* of the point with respect to the camera. This word is derived from the Greek word : $\chi\epsilon\iota\rho$ meaning *hand* or *side*. It is well know that cheirality is valuable in determining scene geometry for calibrated cameras. Longuet-Higgins [7] uses it to distinguish between four different possible scene reconstructions from two views. More recently, Robert and Faugeras ([12]) have used it for the construction of convex hulls of three-dimensional point sets. No systematic treatment of cheirality for uncalibrated cameras has previously appeared, however. Investigation of this phenomenon turns out to be quite fruitful, as is seen in the present paper. Cheirality is valuable in distinguishing different point sets in space, especially in allowing projectively equivalent point sets to be distinguished.

Projective transforms have the property of swapping points from the front to the back of the camera. We will say that a transform is cheirality-reversing for a given point if it swaps the point from the front to the back of the camera, or vice-versa. Otherwise it is called cheirality-preserving. The use of the word cheirality differs slightly from the conventional usage in topology where it refers to local spatial orientation. In topology, a cheirality reversing transform is one that reverses orientation, such as a mapping that takes a point set to its mirror image. For instance, knots that are the same as their mirror image are called *amphicheiral* ([5]). It will be seen in this paper that for affine spatial transforms our definition of cheirality-preserving corresponds with the topological definition in that an orientation preserving transformation preserves the front and back of the cameras. For arbitrary projective transforms the two concepts are distinct.

Summary of major results of the paper. In Definition 4.5 a class of projectivities called *quasi-affine transformations* is defined, consisting of those that preserve the convex hull of a set of points of interest. Theorem 5.14 strengthens the result of [1, 3] by showing that a 3-dimensional point set is determined up to quasi-affine transformation by its image in two uncalibrated views. This sharpening of the theorem of [1, 3] results from a consideration of the cheirality of the cameras. This result leads naturally to the concept of a quasi-affine reconstruction of a scene, which is one that differs by at most a quasi-affine transformation from the true geometry. A practical algorithm for computing a quasi-affine reconstruction of a scene seen in two (or more) views is given in section 8.

Consideration of cheirality leads to a necessary and sufficient condition for a set of image correspondences to be derived as projections of points in a real scene. This is discussed in section 6.

In section 7 the concept of quasi-affine transformation is applied to orientation of point sets, explaining why some point sets allow two differently oriented quasi-affine reconstructions from two views, whereas some do not. The relationship of this result to human visual perception of 3D scenes is briefly mentioned, noting that the brain is able to reconstruct differently oriented quasi-affine models of a scene.

Sections 9 and 10 consider the application of cheirality to view synthesis in which a new view of a scene is constructed from a set of given images.

In section 11.1 a quasi-affine invariant is defined – the cheiral sequence. In section 12 an example of computation of the cheiral sequence for 3D point sets shows that it is useful in distinguishing between non-equivalent point sets. This invariant may be seen as formalizing and extending to three dimensions the thesis and paper of Morin [9, 10] on distinguishing planar shapes.

2 Notation

We will consider object space to be the 3-dimensional Euclidean space R^3 and represent points in R^3 as 3-vectors. Similarly, image space is the 2-dimensional Euclidean space R^2 and points are represented as 2-vectors. Euclidean space, R^3 is embedded in a natural way in projective 3-space $\mathcal{P}^3$by the addition of a plane at infinity. Similarly, R^2 may be embedded in the projective 2-space $\mathcal{P}^2$by the addition of a line at infinity. The simplicity of considering projections between $\mathcal{P}^3$and $\mathcal{P}^2$has led many authors to identify $\mathcal{P}^3$and $\mathcal{P}^2$as the object and images space. This sometimes leads one to forget that real points and cameras lie in Euclidean and not in projective space. Where convenient we will consider points in R^2 and R^3 as lying in $\mathcal{P}^2$and $\mathcal{P}^3$respectively, via the natural embedding. However, in this case the line (or plane) at infinity will be considered to be a special distinguished line (or plane).

Vectors will be represented as bold-face lower case letters, such as $\mathbf{x}$. Such a notation represents a column vector. The corresponding row vector will be denoted by $\mathbf{x}^\top$. The notation $\mathbf{x}$ usually denotes a homogeneous 4-vector representing an element in $\mathcal{P}^3$, whereas $\mathbf{u}$ represents a vector in $\mathcal{P}^2$. The notation $\tilde{\mathbf{x}}$ represents a non-homogeneous 3-vector representing an element of R^3. Similarly, $\tilde{\mathbf{u}}$ is a non-homogeneous 2-vector. The notation $\hat{\mathbf{x}}$ represents a vector with final coordinate equal to 1, sometimes meant implicitly to represent the same point as a homogeneous vector $\mathbf{x}$. Similarly $\hat{\mathbf{u}}$ represents a vector of the form $(u, v, 1)^\top$.

The notation $a \doteq b$ means that a and b have the same sign. For instance $a \doteq 1$ has the same meaning as $a > 0$.

3 Projections in $\mathcal{P}^3$

A projection from $\mathcal{P}^3$into $\mathcal{P}^2$is represented by a 3×4 matrix P, whereby a point $\mathbf{x}$ maps to the point $\mathbf{u} = P\mathbf{x}$. It will be assumed that P has rank 3. Since P has 4 columns but rank 3, there is a unique vector $\mathbf{c}$ such that $P\mathbf{c} = (0, 0, 0)^\top$. In other words, the projective transformation is undefined at the point $\mathbf{c}$, since $(0, 0, 0)^\top$ is not a valid homogeneous 3-vector. The point $\mathbf{c}$ will be called the *perspective centre* of the camera. We will assume for the present that the perspective centre is not a point at infinity so we may write $\mathbf{c} = \hat{\mathbf{c}} = \begin{pmatrix} \tilde{\mathbf{c}} \\ 1 \end{pmatrix}$ where $\mathbf{c}$ is the perspective center as a point in R^3.

Now, the camera matrix P may be written in block form as $P = (M \mid \mathbf{v})$ where M is a 3×3 block and $\mathbf{v}$ is a column vector. Then

$$P\hat{\mathbf{c}} = (M \mid \mathbf{v})\begin{pmatrix} \tilde{\mathbf{c}} \\ 1 \end{pmatrix} = M\tilde{\mathbf{c}} + \mathbf{v} = 0 \ ,$$

and so $\mathbf{v} = -M\tilde{\mathbf{c}}$. Now since P has rank 3 and $-M\tilde{\mathbf{c}}$ is a linear combination of the columns of M, it follows that M must have rank 3. In other words, M is non-singular. Summarizing this discussion we have

Proposition 3.1. *If P is a camera transform matrix for a camera with perspective centre not at infinity, then P can be written as $P = (M \mid -M\tilde{\mathbf{c}})$ where M is a non-singular 3×3 matrix and $\tilde{\mathbf{c}}$ represents the perspective centre in R^3.*

There exist points in $\mathcal{P}^3$that are mapped to points at infinity in the image. To find what they are, we suppose that $\mathbf{u} = (u, v, 0)^\top = P\mathbf{x}$. Letting $\mathbf{p}_1{}^\top$, $\mathbf{p}_2{}^\top$ and $\mathbf{p}_3{}^\top$ be the rows of P, we see that $\mathbf{p}_3{}^\top\mathbf{x} = 0$. In other words, a point $\mathbf{x}$ in $\mathcal{P}^3$that maps to a point at infinity in the image must satisfy the equation $\mathbf{x}^\top\mathbf{p}_3 = 0$. Looked at another way, if $\mathbf{p}_3$ is taken as representing a plane in $\mathcal{P}^3$, then it represents the plane consisting of all points mapping to infinity in the image. Since $P\mathbf{c} = 0$, we see in particular that $\mathbf{p}_3{}^\top\mathbf{c} = 0$ and so $\mathbf{c}$ lies on the plane $\mathbf{p}_3$. To summarize this paragraph, the set of points in $\mathcal{P}^3$mapping to a point at infinity in the image is a plane passing through the perspective centre and represented by $\mathbf{p}_3$, where $\mathbf{p}_3{}^\top$ is the last row of P. In conformity with standard terminology, this plane will be called the *principal plane* of the camera.

Restricting now to R^3, consider a point $\mathbf{x}$ in space, not lying on the principal plane. It is projected by the camera with matrix P onto a point $\mathbf{u}$ where $w\hat{\mathbf{u}} = P\hat{\mathbf{x}}$ for some scale factor w. The value of w will vary continuously with $\mathbf{x}$ and the set of points where it vanishes is precisely the principal plane. It follows that on one side of the principal plane $w > 0$ and on the other side, $w < 0$.

In a Euclidean context, the value of w can be given a precise metric interpretation as explained next. The line perpendicular to the principal plane through the perspective centre is called the *principal ray*. In general, the normal vector to a plane $(q, r, s, t)^\top$ is given in non-homogeneous coordinates as the vector $(q, r, s)^\top$. Thus, if $P = (M \mid -M\tilde{\mathbf{c}})$, then the principal ray is represented by the last row of M, denoted $\mathbf{m}_3{}^\top$.

For a point $\mathbf{x}$ in space, we see that

$$\begin{aligned} w\hat{\mathbf{u}} &= P\hat{\mathbf{x}} \\ &= (M \mid -M\tilde{\mathbf{c}})\begin{pmatrix} \tilde{\mathbf{x}} \\ 1 \end{pmatrix} \\ &= M\tilde{\mathbf{x}} - M\tilde{\mathbf{c}} \\ &= M(\tilde{\mathbf{x}} - \tilde{\mathbf{c}}) \ , \end{aligned}$$

and so $w = \mathbf{m}_3{}^\top(\tilde{\mathbf{x}} - \tilde{\mathbf{c}})$ As just remarked, $\mathbf{m}_3$ represents the direction of the principal ray, and $\tilde{\mathbf{x}} - \tilde{\mathbf{c}}$ is the vector from the camera centre to the point $\mathbf{x}$. If P is scaled by multiplication by an appropriate factor so that $||\mathbf{m}_3|| = 1$ then, w is equal to the depth of the point $\mathbf{x}$ from the camera perspective centre in the direction of the principal ray. This metric interpretation of w, though useful in some applications, such as depth recovery ([14]) will not be used further in this paper.

Any real camera can only view points on one side of the principal plane, those points that are "in front of" the camera. Points on the other side will not be visible. In order to distinguish the front of the camera from the back, a convention is necessary.

Definition 3.2. A camera matrix $P = (M \mid \mathbf{v})$ is said to be *normalized* if $\det(M) > 0$. If P is a normalized camera matrix, a point $\mathbf{x}$ is said to lie *in front of* the camera if $P\hat{\mathbf{x}} = w\hat{\mathbf{u}}$ with $w > 0$. Points $\mathbf{x}$ for which $w < 0$ are said to be *behind* the camera.

Any camera matrix may be normalized by multiplying it by -1 if necessary. The selection of which side of the camera is the front is simply a convention, consistent with the assumption that for a camera with matrix $(I \mid 0)$, points with positive z-coordinate lie in front of the camera. This is the usual convention in computer vision literature, used for instance in [7].

To avoid having always to deal with normalized camera matrices, we define the following parameter χ.

Definition 3.3. Suppose a point $\mathbf{x} = (x, y, z, t)^\top$ maps to an image point $\mathbf{u} = (u, v, w)^\top$ by a camera with matrix $P = (M \mid \mathbf{v})$. Thus, let $(u, v, w)^\top = P(x, y, z, t)^\top$. We define

$$\chi(\mathbf{x}; P) = (\det M)^{1/3} t/w$$

□

Note that the value of χ is unchanged if the point $\mathbf{x}$ is multiplied by a non-zero scale, since the value of w is multiplied by the same scale. Similarly, if P is multiplied by a constant scale k, then both $\det M^{1/3}$ and w are multiplied by k, and the value of χ is unchanged. Thus, $\chi(\mathbf{x}; P)$ is independent of the particular homogeneous representation of $\mathbf{x}$ and P. If P is normalized and $t = 1$ so that $\mathbf{x} = \hat{\mathbf{x}}$, then $\chi(\mathbf{x}; P) \doteq w$. Thus, corresponding to Definition 3.2 we have

Proposition 3.4. *The point $\mathbf{x}$ lies in front of the camera P if and only if $\chi(\mathbf{x}; P) > 0$.*

In fact, χ is positive for points in front of the camera, negative for points behind the camera, zero on the plane at infinity and infinite on the principal plane of the camera. If the camera centre or the point $\mathbf{x}$ is at infinity, then χ is still defined but is equal to zero. In this case, it is not well defined whether the point lies behind or in front of the camera.

In general, we will only be concerned with the sign of χ and not its magnitude. We may then write $\chi(\mathbf{x}; P) \doteq t \det M/w$ (remember that the symbol $\doteq$ indicates equality of sign). The quantity $\mathrm{sign}(\chi(\mathbf{x}; P))$ will be referred to as the *cheirality* of the point $\mathbf{x}$ with respect to the camera 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.

Note on the figures. In the figures included in this paper, a non-standard representation of cameras is used. A camera is denoted by a line representing its principal plane, along with an arrow pointing in the direction of the front of the camera. The tail of the arrow lies at the centre of projection, on the principal plane. Generally, the figures contain one or two cameras. The diagrams may be thought of as representing the projection of R^3 along the direction of the common line of intersection of the two cameras' principal planes. Thus, each principal plane projects to a line, and their line of intersection projects to a point.

4 Quasi-Affine Transformations

A subset B of R^n is called convex if the line joining any two points in B also lies entirely within B. The convex hull of B, denoted $\bar{B}$ is the smallest convex set containing B. We denote by L_∞ the $(n-1)$-dimensional subspace (line, plane, etc) at infinity in $\mathcal{P}^n$. For simplicity, we will refer to it as the plane at infinity, except where we are specifically considering $\mathcal{P}^2$. The inverse image of L_∞ under a projective transformation h is denoted $\boldsymbol{\pi}_\infty = h^{-1}(L_\infty)$.

Definition 4.5. Let B be a subset of R^n and let h be a projectivity of $\mathcal{P}^n$. The projectivity h is said to be "quasi-affine" with respect to the set B if $h^{-1}(L_\infty)$ does not meet $\bar{B}$, where L_∞ is the plane at infinity.

A projectivity that is quasi-affine with respect to B is precisely one that preserves the convex hull of B (as will be seen later).

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 $\{\mathbf{x}_i\}$.

An alternative characterization of quasi-affine transformations is given in the following theorem.

Theorem 4.6. *A projectivity $h : \mathcal{P}^n \rightarrow \mathcal{P}^n$ represented by a matrix H is quasi-affine with respect to a set $B = \{\mathbf{x}_i\} \subset R^n - h^{-1}(L_\infty)$ if an only if there exist constants w_i,* all of the same sign*, such that $H\hat{\mathbf{x}}_i = w_i\hat{\mathbf{x}}'_i$*

Proof. To prove the forward implication, we assume that h is a quasi-affine transformation. By definition, constants w_i exist such that $H\hat{\mathbf{x}}_i = w_i\hat{\mathbf{x}}'_i$. What needs proof is that they all have the same sign. The value of w in the mapping $w\hat{\mathbf{x}}' = H\hat{\mathbf{x}}$ is a continuous function of the point $\mathbf{x}$. If $w_i > 0$ for some point $\mathbf{x}_i$, and $w_j < 0$ for another point $\mathbf{x}_j$, then there must exist some point $\mathbf{x}_\infty$ on the line segment joining $\mathbf{x}_i$ to $\mathbf{x}_j$ for which $w = 0$. This means that $\mathbf{x}_\infty$ lies in $\bar{B}$, but $h(\mathbf{x}_\infty)$ lies on the line at infinity, contrary to hypothesis.

Next, to prove the converse, we assume that there exist such constants w_i all of the same sign. We need to show that $h^{-1}(L_\infty)$ does not meet $\bar{B}$. Let S be the subset of R^n consisting of all points $\mathbf{x}$ satisfying the condition $H\hat{\mathbf{x}} = w\hat{\mathbf{x}}'$ such that w has the same sign as all w_i. The set S contains B and it is clear that $S \cap h^{-1}(L_\infty) = \emptyset$. All that remains to show is that S is convex, for then S must contain the convex hull of B. If $\mathbf{x}_i$ and $\mathbf{x}_j$ are points in S with corresponding constants w_i and w_j, then any intermediate point $\mathbf{x}$ between $\mathbf{x}_i$ and $\mathbf{x}_j$ must have w value intermediate between w_i and w_j. To see this, consider a point $\hat{\mathbf{x}} = \alpha\hat{\mathbf{x}}_i + (1-\alpha)\hat{\mathbf{x}}_j$ where $0 \leq \alpha \leq 1$. This point lies between $\mathbf{x}_i$ and $\mathbf{x}_j$. Denote by $\mathbf{h}_4{}^\top$ the last row of H. Then,

$$\begin{aligned} w &= \mathbf{h}_4{}^\top\hat{\mathbf{x}} \\ &= \mathbf{h}_4{}^\top(\alpha\hat{\mathbf{x}}_i + (1-\alpha)\hat{\mathbf{x}}_j) \\ &= \alpha\mathbf{h}_4{}^\top\hat{\mathbf{x}}_i + (1-\alpha)\mathbf{h}_4{}^\top\hat{\mathbf{x}}_j \\ &= \alpha w_i + (1-\alpha)w_j \end{aligned}$$

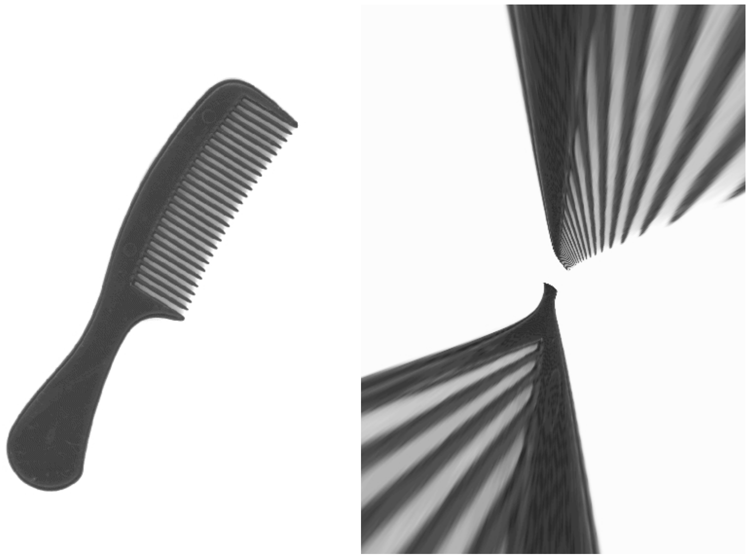


Figure 1: *Picture of a comb and a non-quasi-affine resampling of the comb*

which lies between w_i and w_j as claimed. Consequently, the value of w must have the same sign as w_i and w_j, and so $\mathbf{x}$ lies in S also. This completes the proof. □

This theorem gives an effective method of identifying quasi-affine mappings. The question remains whether quasi-affine mappings form a useful class. This question will be answered by the following theorem.

Theorem 4.7. *If B is a point set in a plane (the "object plane") in 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 proven 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 $L = h^{-1}(L_\infty)$, where L_∞ is the line at infinity in the image plane. Therefore, h is a quasi-affine with respect to B. □

As an example to illustrate the difference between projective and quasi affine mapping, consider Fig. 1. This figure shows an image of a comb and the image resampled according to a projective mapping. Most people will agree that the resampled image is unlike any view of a comb seen by camera or human eye. Nevertheless, the two images are projectively equivalent and will have the same projective invariants. The projective mapping is **not**, however, quasi-affine with respect to the comb.

Note that if points $\mathbf{u}_i$ are visible in an image, then the corresponding object points must lie in front of the camera. Applying Theorem 4.7 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, $\{\mathbf{u}_i\}$ and $\{\mathbf{u}'_i\}$ being correponding points in the two images, then there is a quasi-affine mapping (with respect to the $\mathbf{u}_i$) mapping each $\mathbf{u}_i$ to $\mathbf{u}'_i$, and so Theorem 4.6 applies, yielding the following proposition.

Proposition 4.8. *If $\{\mathbf{u}_i\}$ and $\{\mathbf{u}'_i\}$ are corresponding points in two views of a set of object points $\{\mathbf{x}_i\}$ lying in a plane, then there is a matrix H representing a planar projectivity such that $H\hat{\mathbf{u}}_i = w_i\hat{\mathbf{u}}'_i$ and all w_i have the same sign.*

This fact was previously discovered and exploited by Andrew Zisserman and Charles Rothwell (private communication) and served as a starting point for the current investigation. They derived this result using the methods of [14].

5 Three dimensional point sets

We now consider three-dimensional point sets seen in a pair of images. The 3D locations of the points will be assumed unknown, but image point matches $\mathbf{u}_i \leftrightarrow \mathbf{u}'_i$ will be known. It will be assumed that sufficiently many point matches knowf for the matrix F to be determined unambiguously, that is at least 8 points ([7]). Under these conditions as shown in [3] and [1] it is possible to determine the location of points $\mathbf{x}_i$ and cameras P and P' such that $\mathbf{u}_i = P\mathbf{x}_i$ and $\mathbf{u}'_i = P'\mathbf{x}_i$, and furthermore, the choice is unique up to projectivity of $\mathcal{P}^3$. Recalling the definition of χ (definition 3.3) and Proposition 3.4, if $\chi(\mathbf{x}_i; P)$ and $\chi(\mathbf{x}_i; P')$ are both positive, then the point $\mathbf{x}_i$ lies in front of both cameras, and maps to points $\mathbf{u}_i$ and $\mathbf{u}'_i$ in the two images. Normally, this will not be the case. It is possible, however, that another choice of P, P' and $\mathbf{x}_i$ exists with the desired property.

We introduce some new terminology. A triplet $(F, \{\mathbf{u}_i\}, \{\mathbf{u}'_i\})$ is called an *epipolar configuration* if F is a rank 2 matrix satisfying the epipolar constraint equation $\mathbf{u}'^{\top}_i F\mathbf{u}_i = 0$ for all i. A *weak realization* of $(F, \{\mathbf{u}_i\}, \{\mathbf{u}'_i\})$ is a triplet $(P, P', \{\mathbf{x}_i\})$, where P and P' are a choice of camera matrices corresponding to the fundamental matrix F and the points $\{\mathbf{x}_i\}$ are object points satisfying the equations $\mathbf{u}_i = P\mathbf{x}_i$ and $\mathbf{u}'_i = P'\mathbf{x}_i$ for each i. A *strong realization* is such a triplet satisfying the additional condition that $\chi(\mathbf{x}_i; P) > 0$ and $\chi(\mathbf{x}_i; P) > 0$ for all i. This condition implies that the points and the camera centres are at finite points. The triplet $(F, \{\mathbf{u}_i\}, \{\mathbf{u}'_i\})$ is called a *feasible configuration* if a strong realization exists. The purpose of considering epipolar configurations, rather than simply a set of point correspondences $\mathbf{u}_i \leftrightarrow \mathbf{u}'_i$ is to avoid the problem of having insufficiently many points, or critical configurations of points that make unique determination of the fundamental matrix impossible. The fundamental matrix will be assumed known. Another common terminology that expresses the same thing is that the cameras are "weakly calibrated".

At this point, it is desirable to derive a slightly different form of the definition of the function χ defined in Definition 3.3. In this definition, and henceforth, we allow the possibility that the camera is located at infinity. Let P be a camera matrix. The centre of P is the unique point $\mathbf{c}$ such that $P\mathbf{c} = 0$. One can write an explicit formula for $\mathbf{c}$ as follows.

Definition 5.9. Given a camera matrix P, we define $\mathbf{c}_P{}^{\top}$ to be the vector (c_1, c_2, c_3, c_4), where

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

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

For convenience of typesetting, we introduce the notation $(P/\mathbf{v}^{\top})$ to represent a 4×4 matrix made up of a 3×4 camera matrix P augmented with an final row $\mathbf{v}^{\top}$. Definition 5.9 leads to a simple formula for $\det(P/\mathbf{v}^{\top})$. Cofactor expansion of the determinant along the last row gives $\det(P/\mathbf{v}^{\top}) = \mathbf{v}^{\top}\mathbf{c}_P$ for any row vector $\mathbf{v}^{\top}$. As a special case, if $\mathbf{p}_i{}^{\top}$ is the i-th row of P, then

$$\mathbf{p}_i{}^{\top}\mathbf{c}_P = \det(P/\mathbf{p}_i{}^{\top}) = 0$$

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

Note that submatrix $\hat{P}^{(4)}$ is the same as matrix M in the decomposition $P = (M \mid \mathbf{v})$, and so $\det M = c_4$. This allows us to reformulate the definition of χ as given in Definition 3.3, as follows.

$$\chi(\mathbf{x}; P) \doteq (\mathbf{e}_4{}^{\top}\mathbf{x})(\mathbf{e}_4{}^{\top}\mathbf{c})/w \tag{1}$$

where $\mathbf{c} = \mathbf{c}_P$ as defined in Definition 5.9, and $\mathbf{e}_4{}^\top$ 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 $\mathbf{e}_4{}^\top\mathbf{x} = 0$.

5.1 Effect of Transformations on Cheirality

We now consider a projective transformation represented by matrix H. Writing $P' = PH^{-1}$ and $\mathbf{x}' = H\mathbf{x}$ one sees that $P\mathbf{x} = P'\mathbf{x}'$. So if $\mathbf{u} = P\mathbf{x}$ then $\mathbf{u} = P'\mathbf{x}'$. Thus, 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 $H\mathbf{x}$ and the camera matrix is transformed to PH^{-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 $\mathbf{c}_P$ when P is transformed to PH^{-1}. To answer that question, consider an arbitrary 4-vector $\mathbf{v}$. We see that

$$\mathbf{v}^\top H^{-1}\mathbf{c}_{PH^{-1}} = \det(PH^{-1}/\mathbf{v}^\top H^{-1}) = \det(P/\mathbf{v}^\top)\det H^{-1} = \mathbf{v}^\top\mathbf{c}_P \det H^{-1} \ .$$

Since this is true for all vectors $\mathbf{v}$, it follows that $H^{-1}\mathbf{c}_{PH^{-1}} = \mathbf{c}_P \det H^{-1}$, or

$$\mathbf{c}_{PH^{-1}} = H\mathbf{c}_P \det H^{-1} \tag{2}$$

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

Now, applying (2) to (1) gives

$$\begin{aligned}\chi(H\mathbf{x}; PH^{-1}) &\doteq (\mathbf{e}_4{}^\top H\mathbf{x})(\mathbf{e}_4{}^\top\mathbf{c}_{PH^{-1}})/w \\ &\doteq (\mathbf{e}_4{}^\top H\mathbf{x})(\mathbf{e}_4{}^\top H\mathbf{c})\det H^{-1}/w\end{aligned}$$

where $\mathbf{c} = \mathbf{c}_P$. Finally, denoting the fourth row of the transformation matrix H by $\mathbf{h}_4{}^\top$, and $\mathrm{sign}(\det H)$ by δ, we obtain

$$\chi(H\mathbf{x}; PH^{-1}) \doteq \delta(\mathbf{h}_4{}^\top\mathbf{x})(\mathbf{h}_4{}^\top\mathbf{c})/w \ . \tag{3}$$

This equation will be used extensively. Note that it may be considered to be a generalization of (1) as will now be explained. A point $\mathbf{x}$ is mapped to the plane at infinity by H if and only if $\mathbf{h}_4{}^\top\mathbf{x} = 0$. Interpreting $\mathbf{h}_4$ as the coordinates of a plane, this condition means that $\mathbf{h}_4$ represents the plane mapped to infinity by H. The factor $\delta \doteq \det H^{-1}$ represents the change of spatial orientation effected by the transformation H, in that H is orientation-preserving if $\det H > 0$ and orientation-reversing if $\det H < 0$. This point will be explained more fully in section 7. Thus, the terms in (3) may be interpreted as follows : $\mathbf{x}$ are the point coordinates; $\mathbf{c}$ are the coordinates of the camera centre, as in Definition 5.9; $\mathbf{h}_4$ are the coordinates of the plane at infinity and δ is the spatial orientation. Compare this with (1) in which $\mathbf{e}_4$ represents the plane at infinity.

We now consider the effect of different transformations on the cheirality of points with respect to a camera. An affine transformation is one represented by a matrix H for which $\mathbf{h}_4{}^\top = \mathbf{e}_4{}^\top = (0, 0, 0, 1)$. The effect of an affine transformation may now be described.

Proposition 5.10. *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. From (1) and (3) we see that $\chi(\mathbf{x}; P) \doteq \chi(H\mathbf{x}; PH^{-1})\det H$ from which the result follows. □

We now determine how an arbitrary projective transformation affects cheirality.

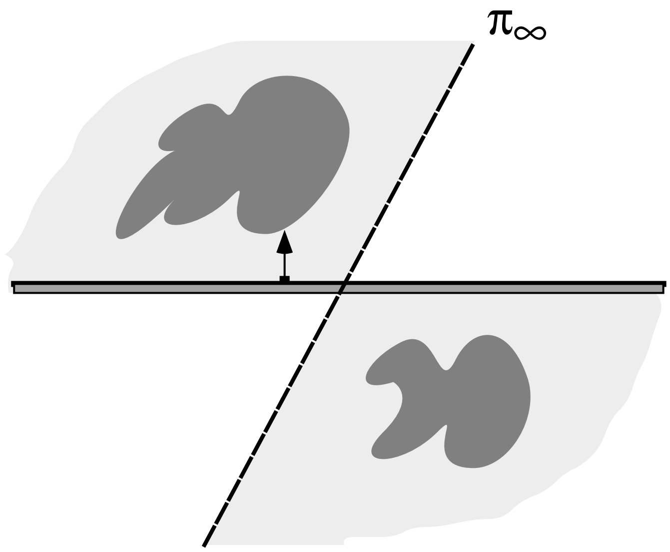


Figure 2: **Effect of a projective transform with positive determinant.** *The principal plane of the camera and the plane $\boldsymbol{\pi}_\infty$ divide R^3 into four segments. One pair of opposite segments (shown shaded) are transformed to points in front of the camera. The opposite pair of segments are transformed to points behind the camera. In the local neighbourhood of the camera centre the front and back of the camera are preserved. This consideration determines which pair of segments become the front of the camera. Thus the two dark shaded sets of points lie in front of the camera after transformation. For a transform with negative determinant the opposite pair of segments become the front of the camera.*

Proposition 5.11. *Let H represent a projective transformation with positive determinant, and let $\boldsymbol{\pi}_\infty$ be the plane in space mapped to infinity by H. The cheirality of a point $\mathbf{x}$ is preserved by H if and only if $\mathbf{x}$ lies on the same side of the plane $\boldsymbol{\pi}_\infty$ as the camera centre.*

Proof. Since $\det H > 0$, we see from (1) and (3) that $\chi(\mathbf{x}; P) \doteq \chi(H\mathbf{x}; PH^{-1})$ if and only if $(\mathbf{h}_4{}^\top\mathbf{x})(\mathbf{h}_4{}^\top\mathbf{c}) \doteq (\mathbf{e}_4{}^\top\mathbf{x})(\mathbf{e}_4{}^\top\mathbf{c})$. Suppose the point $\mathbf{x}$ and the camera P are located at finite points so that the cheirality is well defined, and let them be scaled so that $\mathbf{x} = \hat{\mathbf{x}}$ and $\mathbf{c} = \hat{\mathbf{c}}$. In this case, $(\mathbf{e}_4{}^\top\mathbf{x})(\mathbf{e}_4{}^\top\mathbf{c}) = 1$ and we see that cheirality is preserved, if and only if $(\mathbf{h}_4{}^\top\hat{\mathbf{x}})(\mathbf{h}_4{}^\top\hat{\mathbf{c}}) \doteq 1$, or otherwise expressed $\mathbf{h}_4{}^\top\hat{\mathbf{x}} \doteq \mathbf{h}_4{}^\top\hat{\mathbf{c}}$. Since $\mathbf{h}_4$ represents the plane $\boldsymbol{\pi}_\infty$, 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 H, if and only if it lies on the same side of the plane $\boldsymbol{\pi}_\infty$ as the camera centre. □

Points $\mathbf{x}$ close to the camera centre will lie on the same side of $\boldsymbol{\pi}_\infty$ as the camera centre, and hence, their cheirality will be preserved. Thus, Proposition 5.11 implies that cheirality is preserved in a local neighbourhood of the camera centre. This is illustrated in Fig 2.

5.2 Quasi-affine invariance of strong realizations

For planar object sets, Theorem 4.7 established the existence of a quasi-affine mapping between the object plane and the image plane. For *non-planar* objects seen in two views, strong realizations of the epipolar configuration take the rôle played by sets of image points in the two dimensional case.

Theorem 5.12. *Let $(F, \{\mathbf{u}_i\}, \{\mathbf{u}'_i\})$ be an epipolar configuration and let $(P, P', \{\mathbf{x}_i\})$ and $(\bar{P}, \bar{P}', \{\bar{\mathbf{x}}_i\})$ be two separate strong realizations of the configuration. Then the projectivity h mapping each point $\mathbf{x}_i$ to $\bar{\mathbf{x}}_i$ is quasi-affine.*

Proof. If the projectivity is not quasi-affine, then there are points on both sides of $\boldsymbol{\pi}_\infty = h^{-1}(L_\infty)$. Since h preserves the cheirality of points lying on only one side of $\boldsymbol{\pi}_\infty$ it follows that h does not preserve the cheirality of all points, Therefore at least one of the realizations can not be a strong realization, and so the hypothesis that h is not quasi-affine is not tenable. □

The particular case where one of the two realizations is the "correct" realization is of interest. It is the analogue in three dimensions of Proposition 4.7.

Corollary 5.13. *If $\{\mathbf{x}_i\}$ are points in R^3, image coordinates $\{\mathbf{u}_i\}$ and $\{\mathbf{u}'_i\}$ are corresponding image points in two uncalibrated views from which the fundamental matrix F is determined uniquely, and $(P, P', \{\bar{\mathbf{x}}_i\})$ is a strong realization of the triple $(F, \{\mathbf{u}_i\}, \{\mathbf{u}'_i\})$, then there is a quasi-affine mapping taking each $\mathbf{x}_i$ to $\bar{\mathbf{x}}_i$.*

From this corollary, we can deduce one of the main results of this paper.

Theorem 5.14. *Let $(P, P', \{\mathbf{x}_i\})$ and $(\bar{P}, \bar{P}', \{\bar{\mathbf{x}}_i\})$ be two different reconstructions of 3D scene geometry derived as strong realizations of possibly different epipolar configurations corresponding to possibly different pairs of images of a 3D point set. Then there is a quasi-affine transformation mapping each point $\mathbf{x}_i$ to $\bar{\mathbf{x}}_i$.*

What this theorem is saying is that if a point set in R^3 is reconstructed as a strong realization from two separate pairs of views, then the two results are the same up to a quasi-affine transformation.

Proof. By corollary 5.13 there exist quasi-affine transformations mapping each of the sets of reconstructed points $\{\mathbf{x}_i\}$ and $\{\bar{\mathbf{x}}_i\}$ to the actual 3D locations of the points. The result follows by composing one of these projectivities with the inverse of the other. □

6 When are a Set of Image Correspondences Realizable ?

Given a set of image correspondences $\mathbf{u}_i \leftrightarrow \mathbf{u}'_i$ one may ask under what conditions these correspondences may arise from projection of points in a real scene into the two images. A well known constraint is the epipolar constraint $\mathbf{u}'^{\top}_i F \mathbf{u}_i = 0$ for some rank-2 matrix, the fundamental matrix. It is shown here that that condition is not sufficient, and a necessary and sufficient condition will be given.

As usual, we avoid the problem of critical point configurations, or insufficiently many point correspondences by assuming that the images are "weakly calibrated" meaning that the fundamental matrix is given. In the terminology already introduced, we assume that we have an epipolar configuration $(F, \{\mathbf{u}_i\}, \{\mathbf{u}'_i\})$. It has been shown in [3, 1] that a realization $(P, P', \{\mathbf{x}_i\})$ of this configuration exists, and that further, all realizations may be reached from this realization by applying a projective transformation.

Given a realization $(P, P', \{\mathbf{x}_i\})$ we write $P\mathbf{x}_i = w_i\hat{\mathbf{u}}_i$ and $P'\mathbf{x}_i = w'_i\hat{\mathbf{u}}'_i$. Suppose that there is a transformation H that transforms this to a strong realization. This means that $\chi(H\mathbf{x}_i; PH^{-1}) > 0$ and $\chi(H\mathbf{x}_i; P'H^{-1}) > 0$ for all i, from which it follows that $\chi(H\mathbf{x}_i; PH^{-1}) \doteq \chi(H\mathbf{x}_i; P'H^{-1})$ for all i. Substituting the formula (3) gives

$$(\mathbf{h}_4{}^\top \mathbf{x}_i)(\mathbf{h}_4{}^\top \mathbf{c})\delta/w_i \doteq (\mathbf{h}_4{}^\top \mathbf{x}_i)(\mathbf{h}_4{}^\top \mathbf{c}')\delta/w'_i \ .$$

Cancelling common terms from both sides gives

$$(\mathbf{h}_4{}^\top \mathbf{c})/w_i \doteq (\mathbf{h}_4{}^\top \mathbf{c}')/w'_i \ .$$

Now $(\mathbf{h}_4{}^\top \mathbf{c})$ and $(\mathbf{h}_4{}^\top \mathbf{c}')$ must be non-zero, since $\chi(H\mathbf{x}_i; PH^{-1})$ and $\chi(H\mathbf{x}_i; P'H^{-1})$ are non-zero. Rearranging terms leads to $w_i w'_i \doteq (\mathbf{h}_4{}^\top \mathbf{c})(\mathbf{h}_4{}^\top \mathbf{c}')$. Since the right side does not depend on i, this means that $w_i w'_i$ has constant sign for all i, which proves the following proposition.

Proposition 6.15. *Let $(P, P', \{\mathbf{x}_i\})$ be a realization of a feasible epipolar configuration. Write $P\mathbf{x}_i = w_i\hat{\mathbf{u}}_i$ and $P'\mathbf{x}_i = w'_i\hat{\mathbf{u}}'_i$. Then $w_i w'_i$ has the same sign for all i.*

Proposition 6.15 has a geometric interpretation as follows. The principal plane of a camera separates R^3 into two regions. For points on one side of the principal plane $P\mathbf{x}_i = w_i\hat{\mathbf{u}}_i$ with $w_i > 0$, whereas on the other side, $w_i < 0$. The two principal planes divide up R^3 into four quadrants. The condition that $\text{sign}(w_i w'_i)$ is constant corresponds to the geometric condition that the points $\mathbf{x}_i$ all lie in a pair of opposite quadrants.

A Sufficient Condition Proposition 6.15 gives a necessary condition for an epipolar configuration to be feasible. It will next be shown that this condition is also sufficient. This will be done by explicitly showing how the weak realization may be transformed to a strong realization. To ensure that this is possible, we need two extra conditions.

Condition 6.16.

1. The image coordinates of the points $\mathbf{x}_i$ as seen by two cameras are bounded.
2. At least one of the camera centres is not a limit point of the point set X.

Since image coordinates are unchanged under transformation, the first condition is independent of the particular weak realization considered. The second condition concerning limit points is unchanged under continuous transformations. Since the transformations we consider are continuous in a neighbourhood of the camera centres, this condition is also independent of the particular weak realization considered. In any reasonable imaging situation, both these conditions will hold. For finite point sets the two conditions are trivially satisfied. For infinite point sets, the image coordinates of the points will still be limited by the extent of the image, so the first condition will hold. For a topologically closed point set, the second condition will hold, since a point that coincides with the camera centre can not be imaged. In general, for arbitrary point sets, it will not normally be the case that the points can lie arbitrarily close to the camera centre.

This condition may be illustrated graphically as in Fig 3.

Now, we proceed to transform an arbitrary weak realization into a strong realization. We proceed in steps. As a preliminary step, we need to ensure that neither of the two camera centres lies on the plane at infinity. If this were to occur, then we would choose a new weak realization for which the camera centres do not lie on the plane at infinity.

The principal planes of the two cameras must now meet in a line in space. Consider a plane $\boldsymbol{\pi}_\infty$ containing that line, but not equal to either of the two principal planes. This plane will be contained in two opposite quadrants of R^3, except where it meets the two principal planes. Let this plane also be chosen so that it passes through the two quadrants of space that do not contain any of the points $\mathbf{x}_i$. This situation is shown in Fig 4. In this case the plane $\boldsymbol{\pi}_\infty$ separates the two point sets X_+ and X_- lying in opposite quadrants of space. Now consider the effect of a transformation mapping the plane $\boldsymbol{\pi}_\infty$ to infinity. According to Proposition 5.11, the cheirality of one of the two sets X_+ and X_- (with respect to say the first camera) will be reversed and the cheirality of the other will be preserved by this transformation. Since originally X_+ and X_- have opposite cheirality, after the transformation they will have the same cheirality. In other words, the whole set $X = X_+ \cup X_-$ will lie on the same side of the first camera. The same argument holds for the other camera.

In invoking Proposition 5.11, it was assumed that neither of the camera centres lay on the line of intersection of the two principal planes, and hence on the plane $\boldsymbol{\pi}_\infty$ chosen. If this were to occur, then we would choose instead a plane $\boldsymbol{\pi}_\infty$ slightly displaced from this intersection line but still separating the two sets X_+ and X_-. This is possible since conditions 6.16 ensure that the point set X does not approach the line of intersection of the principal planes.

The case where the two principal planes are identical must also be handled specially. In this case, the plane $\boldsymbol{\pi}_\infty$ is chosen slightly displaced from the cameras' common principal plane, and separating X_+ from X_-.

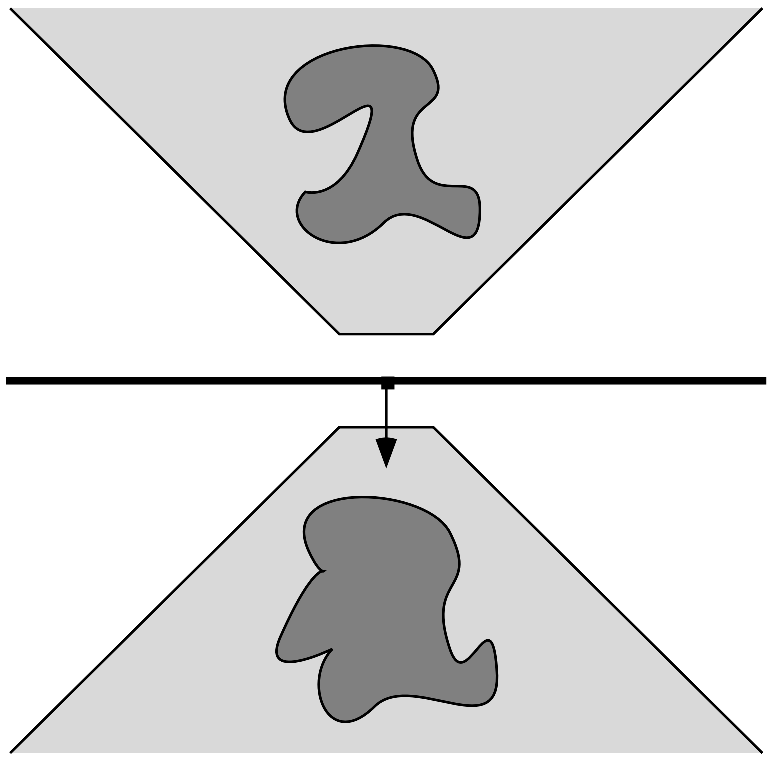


Figure 3: *The point set X (dark shading) must lie inside a truncated cone (dark shading). The cone represents the bounding of the image coordinates. The cone is truncated near the camera centre c since points in X can not lie arbitrarily close to the camera centre. In the general case, points may lie both behind and in front of the camera.*

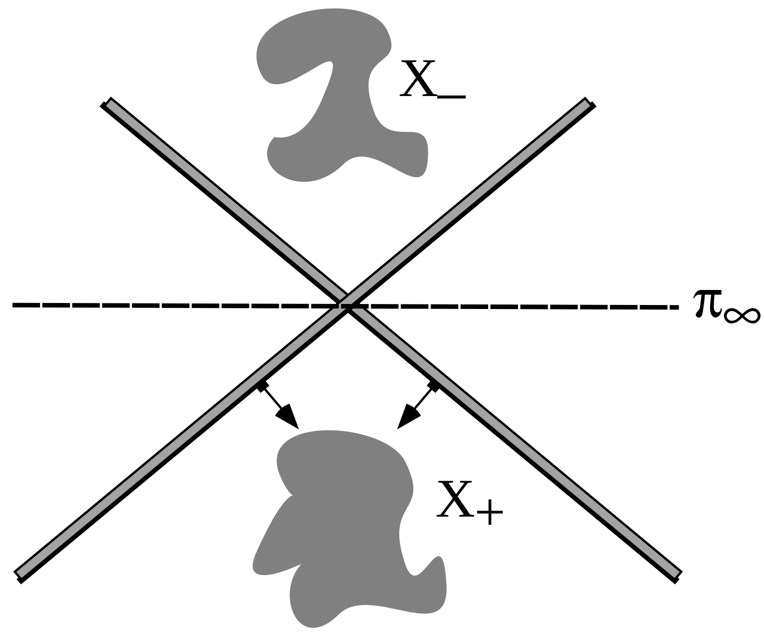


Figure 4: **Step 1 of transformation.** *We choose the plane at infinity to pass through the two quadrants that do not contain the point set. After this transformation, all points will lie on one side of each camera.*

If after this first transformation step, the set X lies in front of both cameras, then we are done. If on the other hand it lies behind both cameras, then applying an affine transformation with negative determinant (for instance $H = \text{diag}(-1, -1, -1, 1)$) will swap the set X to the front of both cameras. There remains the possibility that X lies in front of one camera and behind the other.

To handle this remaining case, we need a further transformation. We wish to find a plane $\boldsymbol{\pi}_\infty$ that separates the two camera centres, but does not separate the point set X. Assuming this is possible, X will then lie on the opposite side of $\boldsymbol{\pi}_\infty$ from one of the camera centres (but not the other). Now we apply a transformation that takes $\boldsymbol{\pi}_\infty$ to infinity. According to Proposition 5.11 the cheirality of X will be reversed with respect to one of the cameras, but not the other. Originally the cheirality of X was opposite with respect to the two cameras, and so after the transformation the cheirality will be the same. This means that X will lie on the same side of both cameras. By applying, if necessary, a cheirality-reversing affine transformation it may be assured that X lies in front of both cameras, and we are done.

It remains to explain how the required plane $\boldsymbol{\pi}_\infty$ is to be found. We suppose that the points X lie in front of the first camera and behind the second camera. We wish to find a plane that separates the two camera centers, but does not separate the point set X. The method for constructing this plane is given in Figures 5, 6 and 7 corresponding to whether the second camera lies behind, in front of, or on the principal plane of the first camera. Details of the construction are given in the captions of the figures.

We can summarize this discussion in the following theorem.

Theorem 6.17. *Let $(F, \{\mathbf{u}_i\}, \{\mathbf{u}'_i\})$ be an epipolar configuration and let $(P, P', \{\mathbf{x}_i\})$ be a realization of that configuration. Suppose that conditions (6.16) are satisfied. Let $P\mathbf{x}_i = w_i\hat{\mathbf{u}}_i$ and $P'\mathbf{x}_i = w'_i\hat{\mathbf{u}}_i$. Then $(F, \{\mathbf{u}_i\}, \{\mathbf{u}'_i\})$ is a feasible configuration if and only if $w_i w'_i$ has the same sign for all i.*

Since an epipolar configuration always possesses a weak realization ([3]), Theorem 6.17 gives a necessary and sufficient condition for an epipolar configuration to be realizable as a three dimensional scene.

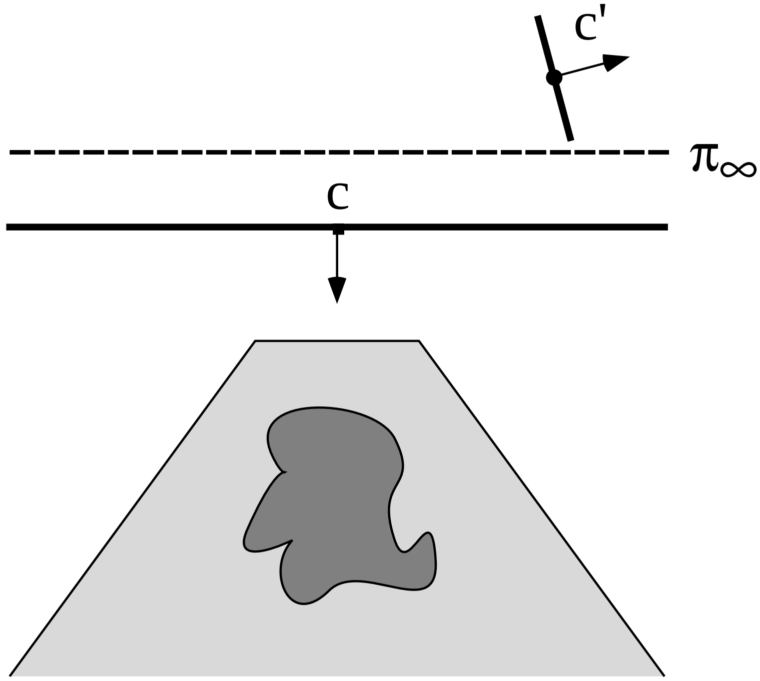


Figure 5: **Second camera behind the first camera** *We can separate the two camera centres* $\mathbf{c}$ *and* $\mathbf{c}'$ *with a plane* $\boldsymbol{\pi}_\infty$ *lying just behind the principal plane of the first camera. Since all the points lie in front of the camera, plane* $\boldsymbol{\pi}_\infty$ *does not separate the point set* X*.*

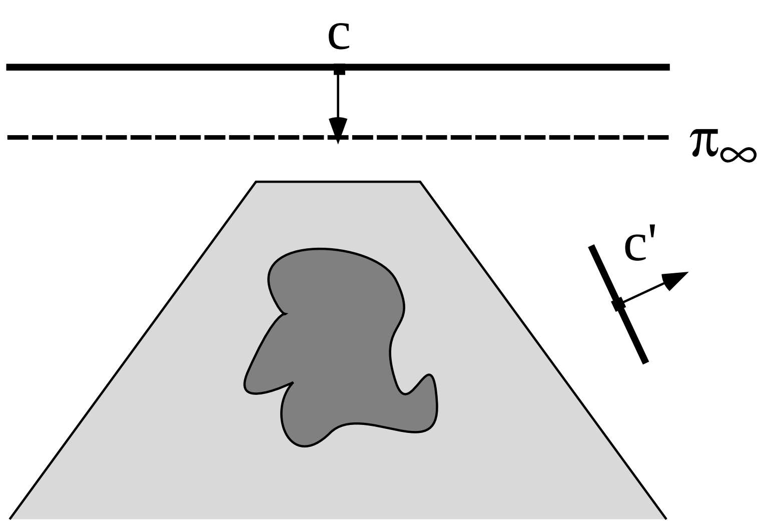


Figure 6: **Second camera in front of the first camera** *We can separate the two camera centres* $\mathbf{c}$ *and* $\mathbf{c}'$ *with a plane* $\boldsymbol{\pi}_\infty$ *lying just in front of the principal plane of the first camera. The point set* X *lies entirely inside the truncated cone (lightly shaded). The plane* $\boldsymbol{\pi}_\infty$ *can be chosen sufficiently close to* $\mathbf{c}$ *so as not to meet this cone. Consequently, it will not separate the point set* X*.*

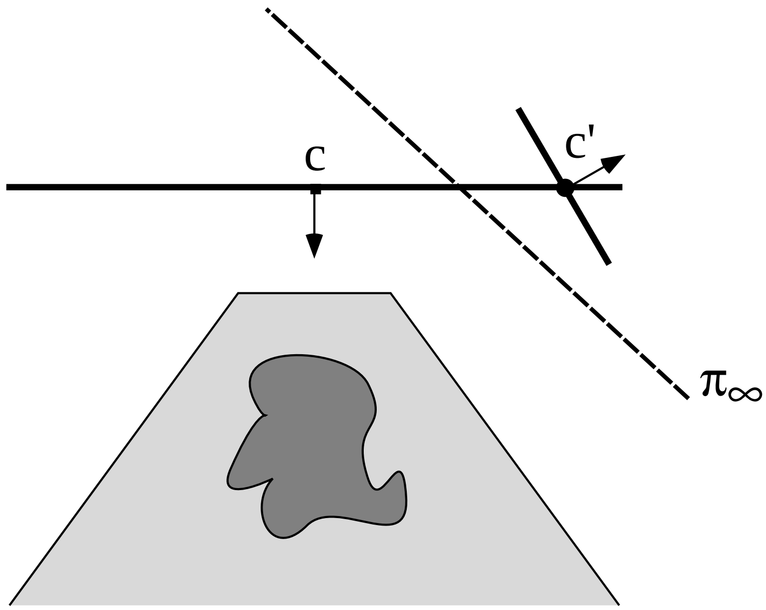


Figure 7: **Second camera lies on the principal plane of the first camera.** *We can separate the two camera centres* $\mathbf{c}$ *and* $\mathbf{c}'$ *with an oblique plane* $\boldsymbol{\pi}_\infty$ *which crosses the principal plane of the first camera. Plane* $\boldsymbol{\pi}_\infty$ *can be chosen so as not to meet the cone containing* X*, and consequently will not separate* X*.*

7 Orientation

We now consider the question of image orientation. A mapping h from 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 R^n with respect to a hyperplane (that is mirror imaging) is an example of an orientation reversing mapping. A projectivity h from $\mathcal{P}^n$ to itself restricts to a mapping from $R^n - h^{-1}(L_\infty)$ to R^n, where L_∞ is the hyperplane (line, plane) at infinity. Consider the case $n = 3$ and let H be a 4×4 matrix representing the projectivity h. We wish to determine at which points $\mathbf{x}$ in $R - h^{-1}(L_\infty)$ the map h is orientation preserving. It may be verified (quite easily using Mathematica [16]) that if $H\hat{\mathbf{x}} = w\hat{\mathbf{x}}'$ and J is the matrix of partial derivatives of h evaluated at $\mathbf{x}$, then $\det(J) = \det(H)/w^4$. This gives the following result.

Proposition 7.18. *A projectivity* h *of* $\mathcal{P}^3$ *represented by a matrix* H *is orientation preserving at any point in* $R^3 - h^{-1}(L_\infty)$ *if and only if* $\det(H) > 0$.

Of course, the concept of orientability may be extended to the whole of $\mathcal{P}^3$, and it may be shown that h is orientation-preserving on the whole of $\mathcal{P}^3$if and only if $\det(H) > 0$. The essential feature here is that as a topological manifold, $\mathcal{P}^3$is orientable. The situation is somewhat different for $\mathcal{P}^2$, which is not orientable as a topological space. In this case, with notation similar to that used above, it may be verified that $\det(J) = \det(H)/w^3$. Therefore, the following proposition is true.

Proposition 7.19. *A projectivity* h *of* $\mathcal{P}^2$*is orientation preserving at a point* $\mathbf{u}$ *in* $R^2 - h^{-1}(L_\infty)$ *if and only if* $w \det(H) > 0$*, where* $H\hat{\mathbf{u}} = w\hat{\mathbf{u}}'$.

This theorem allows us to strengthen the statement of Theorem 4.6 somewhat.

Corollary 7.20. *If* h *is a quasi-affine transformation of* $\mathcal{P}^2$*with respect to a set of points* $\{\mathbf{u}_i\}$ *in* R^2*, then* h *is either orientation-preserving or orientation-reversing at all points* $\mathbf{u}_i$*. Suppose the* *matrix H corresponding to h is normalized to have positive determinant (by possible multiplication by -1) and let $H\hat{\mathbf{u}}_i = w_i \hat{\mathbf{u}}'_i$. Then h is orientation-preserving if and only if $w_i > 0$ for all i.*

An example where Corollary 7.20 applies is in the case where two images of a planar object are taken from the same side of the object plane. In this case, an orientation-preserving quasi-affine projectivity will exist between the two images. Consequently, all the w_i defined with respect to a matrix H will be positive, provided that H is normalized to have positive determinant.

The situation in 3-dimensions is rather more involved and more interesting. Two sets of points $\{\mathbf{x}_i\}$ and $\{\bar{\mathbf{x}}_i\}$ that correspond via a quasi-affine transformation are said to be *oppositely oriented* if the projectivity is orientation-reversing. This definition extends also to two strong realizations $(P, P', \{\mathbf{x}_i\})$ and $(\bar{P}, \bar{P}', \{\bar{\mathbf{x}}_i\})$ of a common epipolar configuration $(F, \{\mathbf{u}_i\}, \{\mathbf{u}'_i\})$, since in view of Theorem 5.12 the point sets are related via a quasi-affine transformation. Whether or not oppositely oriented strong realizations exist depends on the imaging geometry. 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 3-D reconstruction of the scene (a strong realization of the image pair). 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, that in certain circumstances, two oppositely oriented realizations of an image pair exist. It may be surprising to discover that this is not always the case, as is shown in the following theorem.

Theorem 7.21. *Let $(F, \{\mathbf{u}_i\}, \{\mathbf{u}'_i\})$ be an epipolar configuration and let $(P, P', \{\mathbf{x}_i\})$ be a strong realization of $(F, \{\mathbf{u}_i\}, \{\mathbf{u}'_i\})$. There exists a different oppositely oriented strong realization $(\bar{P}, \bar{P}', \{\bar{\mathbf{x}}_i\})$ if and only if there exists a plane in R^3 such that the perspective centres of both cameras P and 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 G be a projective transformation mapping the given plane to infinity and let A be an affine transformation. Suppose further that $\det G > 0$ and $\det A < 0$. Let H be the composition $H = AG$. According to Proposition 5.11 the transformation H is cheirality reversing for the points, since the points are on the opposite side of the plane from the camera centres. According to Proposition 5.10 A is also cheirality reversing, since $\det A < 0$. The composition H must therefore be cheirality preserving, and it transforms the strong configuration to a different strong configuration. Since 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 H be the transformation taking one to the other. Since H is orientation reversing, $\det H < 0$. The mapping 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 H, then according to Propositions 5.11 the points X must lie on the opposite side of the plane $\boldsymbol{\pi}_\infty$ from both camera centres. □

8 The Cheiral Inequalities

Several methods ([1, 3, 8]) have been proposed for computing a projective reconstruction (in our terminology a weak realization) of a scene from a set of point matches. In section 6 a constructive method was given for transforming a weak realization into a strong one. That method was not very suitable for computer computation. Accordingly, in this section a straight-forward algorithm will be given for computing a strong realization of an epipolar configuration. This will be done by transforming a weak realization into a strong realization by finding an appropriate transformation.

We start with a weak realization $(P, P', \{\mathbf{x}_i\})$ of an epipolar configuration. Let $w_i \hat{\mathbf{u}}_i = P\mathbf{x}_i$ and $w'_i \hat{\mathbf{u}}_i = P'\mathbf{x}_i$. We assume that $w_i w'_i$ has the same sign for all i. By multiplying the matrix P by

-1 if necessary, we may ensure that $w_i w'_i > 0$ for all i. Furthermore, by multiplying $\mathbf{x}_i$ by -1 if necessary, we may ensure that $w_i > 0$ and hence $w'_i > 0$ for all i. We will assume that this has been done.

Now, we seek a transformation H that will transform the weak realization to a strong realization. After this transformation, all points will lie in front of both cameras. According to (3) this condition may be written (for camera P)

$$\chi(\mathbf{x}_i; P) \doteq (\mathbf{h}_4{}^\top \mathbf{x}_i)(\mathbf{h}_4{}^\top \mathbf{c})\delta > 0$$

where $\delta = \text{sign}(\det H)$. Similarly, for the other camera, we have

$$\chi(\mathbf{x}_i; P') \doteq (\mathbf{h}_4{}^\top \mathbf{x}_i)(\mathbf{h}_4{}^\top \mathbf{c}')\delta > 0 \ .$$

Since we are free to multiply $\mathbf{h}_4$ by -1 if necessary, we may assume that $(\mathbf{h}_4{}^\top \mathbf{c})\delta > 0$. From this it follows from the first inequality that $\mathbf{h}_4{}^\top \mathbf{x}_i > 0$ for all i. Then, from the second inequality, we have $(\mathbf{h}_4{}^\top \mathbf{c}')\delta > 0$. The total set of inequalities may now be written :

$$\begin{aligned} \mathbf{x}_i{}^\top \mathbf{h}_4 &> 0 \\ \delta \mathbf{c}^\top \mathbf{h}_4 &> 0 \\ \delta \mathbf{c}'^\top \mathbf{h}_4 &> 0 \end{aligned} \qquad (4)$$

These equations (4) 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 $\mathbf{h}_4$. 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 H, first of all we solve the cheiral inequalities to find a value of $\mathbf{h}_4$, either for $\delta = 1$ or $\delta = -1$. The required matrix H is any matrix having $\mathbf{h}_4{}^\top$ as its last row and satisfying the condition $\det H \doteq \delta$. If the last component of $\mathbf{h}_4$ is non-zero, then H can be chosen to have the simple form in which the first three rows are of the form $\pm(I \mid \mathbf{0})$.

Theorem 6.17 guarantees that there will 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 7.

Solving the Cheiral Inequalities Naturally, the cheiral inequalities may be solved using techniques of linear programming. As they stand however, if $\mathbf{h}_4$ is a solution, then so is $\alpha\mathbf{h}_4$ for any positive factor α. In order to restrict the solution domain to be bounded, one may add additional inequalities. For instance, if $\mathbf{h}_4 = (v_1, v_2, v_3, v_4)^\top$, 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 by which each of the inequalites is satisfied. To do this, we introduce one further variable, d. Each of the inequalities $\mathbf{a}^\top \mathbf{h}_4$ of the form (4) for appropriate $\mathbf{a}$ is replaced by an inequality $\mathbf{a}^\top \mathbf{h}_4 > d$. 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 ([11])[1]. If a solution is found for which $d > 0$ then this will be a desired solution.

8.1 Quasi-affine reconstruction

A strong realization of an epipolar configuration is a *quasi affine* reconstruction, since it differs from the true scene by a quasi-affine transformation (Corollary 5.13). Quasi-affine reconstructions

[1]The Simplex algorithm given in [11] 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

of a scene have useful properties such as preservation of complex hull. Furthermore, computing a quasi-affine reconstruction has been used in [6] as a preliminary step towards computing a Euclidean reconstruction of a scene from three views with the same camera. A strong realization of an epipolar reconstruction is a slightly more restrictive than a general quasi-affine reconstruction, however, as will be shown now.

The inequalities (4) are seen to be of two types. The first inequality involves the points (one inequality for each i) and the other two involve the camera centres. One sees that if only the first inequality is satisfied (for all i), but possibly not the ones involving the camera centres, then the solution is less constrained. Instead of all points lying in front of both cameras, all points will lie on the same side of each camera. Thus, if $\delta \mathbf{c}^\top \mathbf{h}_4{}^\top < 0$, then all points will lie behind the first camera, since $\chi(\mathbf{x}_i; P) < 0$. Thus, solving the first inequality for all i is equivalent to the first step of the construction given in section 6. Adding the other two inequalities as well is equivalent to carrying out the second step of section 6. Note now that the transformation carried out in the second step is itself quasi-affine. In fact, referring to Figs 5, 6 and 7 one sees that the plane $\boldsymbol{\pi}_\infty$ does not separate the point set X. Thus, just by solving the first inequality of (4) one obtains a quasi-affine reconstruction of the point set. However, including the two inequalities for the camera locations further constrains the reconstruction to bring it closer to the true Euclidean reconstruction, and so is recommended in most cases.

If one is content with any quasi-affine reconstruction, however, then one can ignore the two last inequalities in (4). An example of when this may be sufficient is when one is computing the cheiral sequence of a set of points, to be described in section 11. In this case, there is a very simple means of solution. The inequalities that we need to solve are of the form $\mathbf{h}_4{}^\top \mathbf{x}_i > 0$ for all i. Recall that we are assuming that each $w_i > 0$ and $w'_i > 0$. This being so, we see that $w_i = \mathbf{p}_3{}^\top \mathbf{x}_i > 0$, where $\mathbf{p}_3{}^\top$ is the third row of the camera matrix P. Thus, we may choose $\mathbf{h}_4 = \mathbf{p}_3$ as the solution to the inequalities. More generally, for any α between 0 and 1, we may choose $\mathbf{h}_4 = \alpha \mathbf{p}_3 + (1-\alpha)\mathbf{p}'_3$, where $\mathbf{p}'_3{}^\top$ is the third row of the other camera matrix P'. This corresponds precisely to the construction of Fig 4.

In the case where the weak realization is carried out in a way such that $P = (I \mid 0)$ (for instance, see the method of [3]), then we have a very easy way to obtain a quasi-affine reconstruction. In this case we choose $\mathbf{h}_4 = \mathbf{p}_3 = (0, 0, 1, 0)^\top$, and

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} .$$

Such an H simply swaps the two last components of any point $\mathbf{x}_i$, and the last two columns of each camera matrix. This gives a very simple way of computing a quasi-affine reconstruction.

1. Carry out a projective reconstruction of the scene for which the first camera has matrix $P = (I \mid 0)$.
2. Swap the last two coordinates of each point $\mathbf{x}_i$ and the last two columns of each camera matrix.

Quasi-affine reconstruction using the cheiral inequalities or the simple algorithm just given extends naturally to reconstruction from several views. There is no analogue of Theorem 6.17 to ensure a solution in the multi-view case, but of course if the input data is derived from real data of a real scene, then a solution will exist.

9 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

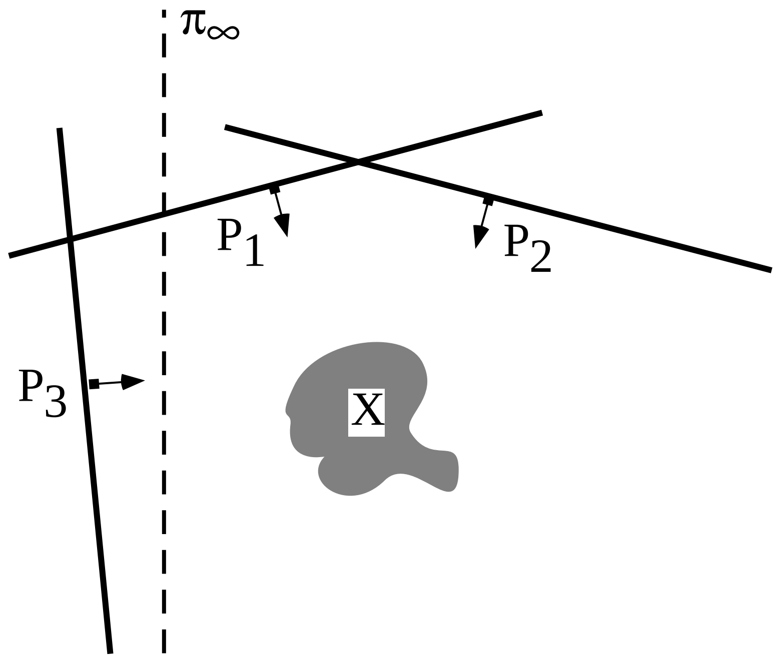


Figure 8: **Visibility.** *In the reconstruction as shown, the point set X lies entirely in front of the first two cameras. Thus, this represents a strong realization of the scene with respect to the first two cameras. As shown, the point set X lies in front of the third camera. However, if the configuration is subjected to a projective transformation so that plane* $\boldsymbol{\pi}_\infty$ *becomes the plane at infinity, then according to Theorem 5.11 the set X will remain in front of the first two cameras, but will be switched to lie behind the third camera. With no way of knowing where the plane at infinity lies, one can not determine whether X lies in front of or behind the third camera.*

reconstruction of the scene from the 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 Fig 8.

Knowledge of a single point known to be visible in the third view serves to break the ambiguity, however, as the following proposition shows. By applying Proposition 6.15 to the first and third views, one obtains the following criterion.

Proposition 9.22. *Let points* $(P^1, P^2, \{\mathbf{x}_i\})$ *be a realization of a set of correspondences* $\mathbf{u}_i^1 \leftrightarrow \mathbf{u}_i^2$*. Let* P^3 *be the camera matrix of a third view and suppose that* $w_j^i \hat{\mathbf{u}}_i = P^i \mathbf{x}_j$ *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 on point visible in the third view. For instance, 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 by solving directly for the matrix P_3 given the correspondences $\mathbf{u}_i^3 = P_3 \mathbf{x}_i$ where points $\mathbf{x}_i$ are the reconstructed points. This may be done by linear means ([15]).

10 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 Fig 9. If a plane exists, separating the camera centres from the point set, then two oppositely oriented reconstructions exist, and one can not determine which points are in front of which.

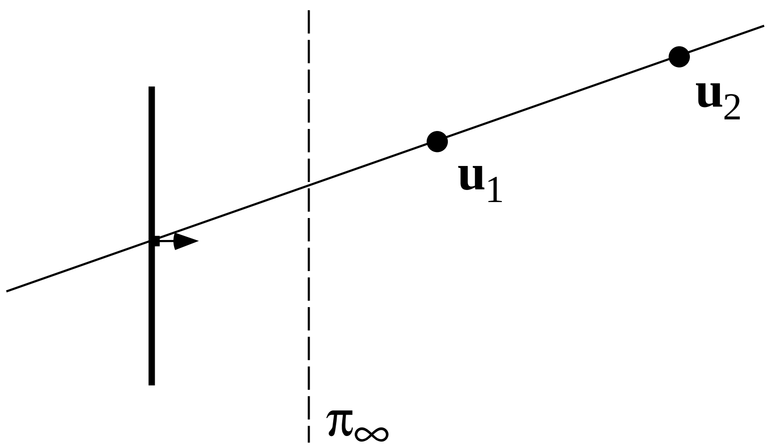


Figure 9: **Which points are in front.** *In the reconstruction shown, point* $\mathbf{u}_1$ *is closer to the third camera than* $\mathbf{u}_2$*. If, however, we apply an orientation-reversing projective transformation that maps the plane* $\boldsymbol{\pi}_\infty$ *to infinity, then the two points will still lie in front of both cameras, but now point* $\mathbf{u}_2$ *will lie closer to the third camera. This is because locally the front and back of the cameras will be reversed by the orientation-reversing transformation. In order to reach* $\mathbf{u}_1$ *from the centre of the third camera, without crossing* $\boldsymbol{\pi}_\infty$ *it is necessary to pass through* $\mathbf{u}_2$ *first.*

The sort of ambiguity shown in Fig 9 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.

The parameter χ defined in Definition 3.3 is used to distinguish the front from the back of the camera in a Euclidean or quasi-affine frame. It is also useful for determining which points lie in front of which, as will be seen now. Recall that χ is zero for points $\mathbf{x}$ on the plane at infinity, infinite for points on the principal plane of the camera, positive for points in front of the camera and negative for points behind the camera. Furthermore, given two points in front of the camera, projecting to the same point in the image , the point with the greater value of χ lies closer to the front of the camera.

The value of χ can be used to parametrize any line in $\mathcal{P}^3$through the camera centre. As one proceeds along the line in the direction of the front of the camera, the value of χ decreases continuously from infinity at the camera centre, through positive values. It reaches zero at the plane at infinity, and continues to decrease through negative values eventually reaching $-\infty$ when the line returns to the camera centre from the rear of the camera. This is illustrated in Fig 10.

Now, if the configuration undergoes a projective transformation H with positive determinant taking the plane $\boldsymbol{\pi}_\infty$ to infinity, then the parameter χ will be replaced by a new parameter χ' defined

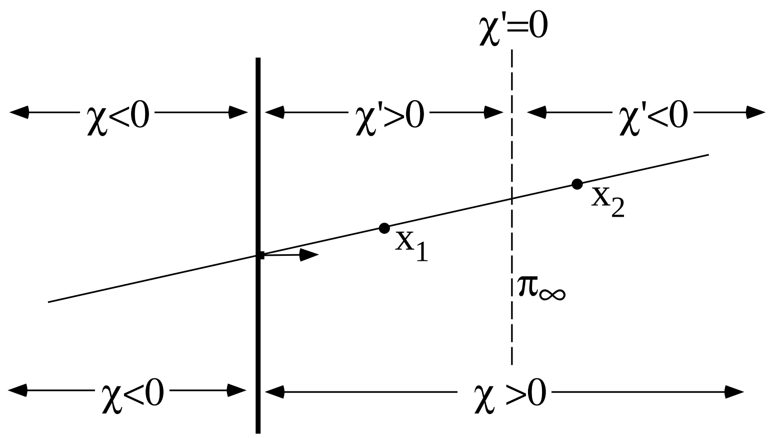


Figure 10: **Preservation of order of points.** *This shows the effect of a transformation with positive determinant taking the plane* $\boldsymbol{\pi}_\infty$ *to infinity. Both* χ *(before the transformation) and* χ' *(after the transformation) decrease monotonically along any ray through the camera centre. We find that* $\chi'(\mathbf{x}_1) > \chi'(\mathbf{x}_2)$ *if and only if* $\chi(\mathbf{x}_1) > \chi(\mathbf{x}_2)$.

by $\chi'(\mathbf{x}) = \chi(H\mathbf{x}; PH^{-1})$. Since the transformation is assumed to have positive determinant, it will preserve the front of the camera locally near the camera centre (by Theorem 5.11). Now, as one proceeds along the line in the same direction as before, the parameter χ' will decrease continuously through positive values from infinity at the camera centre, reaching zero where the line crosses the plane $\boldsymbol{\pi}_\infty$ and then continuing to decrease through negative values until the line returns to the camera centre. Since both χ and χ' decrease monotonically as one proceeds along the line, one sees that 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. In this case $\chi'(\mathbf{x}_1) > \chi'(\mathbf{x}_2)$ if and only if $\chi(\mathbf{x}_1) < \chi(\mathbf{x}_2)$.

If 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 proposition that allows us to determine from an arbitrary projective reconstruction which of two points is closer to the front of the camera.

Proposition 10.23. *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 (1)) 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 theorem 10.23 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 8. If differently oriented strong realizations exist, then as illustrated by Fig 9, 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.

11 3D quasi-affine invariants

One of the important properties of quasi-affine transformations is that they preserve separation by planes as will be explained next.

Proposition 11.24. *Let $\mathbf{x}_0$ and $\mathbf{x}_1$ be two points in space and let $\boldsymbol{\pi}$ be a plane not passing through either of the points. Let h be a quasi-affine transformation with respect to the two points taking $\mathbf{x}_i$ to $\mathbf{x}'_i$ and mapping $\boldsymbol{\pi}$ to a plane $\boldsymbol{\pi}'$. Then $\mathbf{x}_0$ and $\mathbf{x}_1$ lie on the same side of $\boldsymbol{\pi}$ if and only if $\mathbf{x}_0$ and $\mathbf{x}_1$ lie on the same side of $\boldsymbol{\pi}'$.*

Proof. Let $\boldsymbol{\pi}$ be represented by a 4-vector $\mathbf{v}$. The points lie on the same side of $\boldsymbol{\pi}$ if and only if $\mathbf{v}^\top \hat{\mathbf{x}}_0 \doteq \mathbf{v}^\top \hat{\mathbf{x}}_1$. Let H represent the projective transformation. Since H is a quasi-affine we have $\hat{\mathbf{x}}'_i = w_i H \hat{\mathbf{x}}_i$ where w_i has the same sign for $i = 0, 1$. The plane represented by $\mathbf{v}$ is mapped to the plane represented by $\mathbf{v}'$ such that $\mathbf{v}'^\top = \mathbf{v}^\top H^{-1}$. Then $\mathbf{v}'^\top \hat{\mathbf{x}}'_i = (\mathbf{v}^\top H^{-1})(w_i H \hat{\mathbf{x}}_i) = w_i \mathbf{v}^\top \hat{\mathbf{x}}_i$. Since all w_i have the same sign, it follows that $\mathbf{v}^\top \hat{\mathbf{x}}_0 \doteq \mathbf{v}^\top \hat{\mathbf{x}}_1$ if and only if $\mathbf{v}'^\top \hat{\mathbf{x}}'_0 \doteq \mathbf{v}'^\top \hat{\mathbf{x}}'_1$, whence the result. □

Given a point set $\{\mathbf{x}_i\}$ it results from this proposition that the set of planes that do not separate the point set is preserved under quasi-affine transformations. Consequently, the convex hull of a set of points is preserved by quasi-affine transformations as was claimed in section 4.

Proposition 11.24 may be used to define quasi-affine invariant properties of point sets. Let $\boldsymbol{\pi}$ be a plane partitioning the point set into two subsets X_+ and X_-. Applying a quasi-affine mapping the transformed point set will be partitioned into the same two subsets by the transformed plane. Thus for each plane $\boldsymbol{\pi}$ there exists an invariant partitioning of the set of points. If the partitioning plane is defined in terms of the point set itself (such as a plane passing through three specified points), then the resulting partition is invariant under quasi-affine transformation, and may be used for indexing purposes.

11.1 An invariant sequence

A way of finding a better invariant plane than the one defined by three points in the set is now described. We describe this method in general n-dimensional space.

Suppose we are given a set of $N \geq n+2$ points $\{\mathbf{x}_i\}$, $i = 1, \ldots, N$ in R^n. Let $\mathbf{e}_1, \ldots \mathbf{e}_{n+2}$ be points in R^n such that $\{\mathbf{e}_i\}$ form a canonical projective basis for $\mathcal{P}^n$. For $n = 2$, the points $(0,0)^\top$, $(1,0)^\top$, $(0,1)^\top$ and $(1,1)^\top$ will do. Assume that the points $\mathbf{x}_i$ are numbered in such a way that the first $n+2$ of them are in general position (meaning that no $n+1$ of them lie in a codimension 1 hyperplane). In this case, there exists a projectivity g (not in general quasi-affine) such that $g(\mathbf{x}_i) = \mathbf{e}_i$ for $i = 1, \ldots, n+2$. Let $\boldsymbol{\pi}_\infty = g^{-1}(L_\infty)$ be the plane in R^n that is mapped to the plane at infinity by this mapping, g. The invariant partition that we are interested in is the one defined by the plane $\boldsymbol{\pi}_\infty$.

We can define the partition more specifically as follows. Let G be a matrix representing the projective transformation g. For each i we may define points $\mathbf{e}_i$ such that $G\hat{\mathbf{x}}_i = \eta_i \hat{\mathbf{e}}_i$ where $\mathbf{x}'_i$ is the image of $\mathbf{x}_i$ under g. In particular for $i = 1, \ldots, n+2$ the points $\mathbf{e}_i$ are our canonical projective basis. In this way, the set $\{\mathbf{x}_i\}$ is partitioned into those points for which $\eta_i > 0$ and those for which $\eta_i < 0$. In exceptional cases the point $\mathbf{e}_i = g(\mathbf{x}_i)$ may lie on the plane at infinity, in which case we set $\eta_i = 0$. This invariant partitioning is of course dependent on the choice of canonical basis $\{\mathbf{e}_i\}$.

The cheiral sequence. We define $\text{sign}(\eta_i)$ to be $+1$, -1 or 0 according to whether η_i is positive, negative or zero. The sequence of values $\text{sign}(\eta_i)$ for $i = 1, \ldots, N$ is called the *cheiral sequence* of the points $\mathbf{x}_i$. Except for a simultaneous change of sign of all η_i, the cheiral sequence is invariant under quasi-affine transformations.

If desired, it is possible to code the values η_i into a single number according to the formula

$$\chi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \left| \sum_{i=1}^{N} \text{sign}(\eta_i) 3^{i-1} \right| \tag{5}$$

The value $\chi(\mathbf{x}_i)$ is invariant under quasi-affine transformation of the ordered set of points $\mathbf{x}_i$.

We now make the assumption that $\eta_i \neq 0$. In this case the cheiral sequence, along with the projective invariants of the point configuration, constitute a complete quasi-affine invariant. This may be stated as follows.

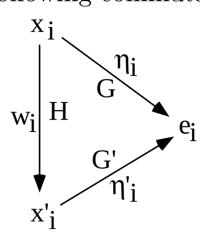
Theorem 11.25. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be a set of points in R^n, where $N \geq n+2$. Suppose that the first $n+2$ of these points form a basis for $\mathcal{P}^n \supset R^n$, so that the cheiral sequence* $\mathrm{sign}(\eta_i)$ *may be defined as above. Suppose further that for each i we have $\eta_i \neq 0$. Let $\mathbf{x}'_1, \ldots, \mathbf{x}'_N$ be another set of points in R^n, projectively equivalent to the points $\{\mathbf{x}_i\}$ via a projective transformation h. Then h is a quasi-affine mapping if and only $\eta_i \doteq \eta'_i \epsilon$ for some constant $\epsilon = \pm 1$.*

Proof. Let points $\mathbf{e}_i$ be defined as in the definition of the cheiral sequence. Further, let g be a projective transform represented by a matrix G and let η_i be defined by the equation $G\hat{\mathbf{x}}_i = \eta_i \hat{\mathbf{e}}_i$. Similarly, we may define projective transformation g' represented by matrix G' and values η'_i such that $G'\hat{\mathbf{x}}'_i = \eta'_i \hat{\mathbf{e}}_i$.

Since the transformation g is defined uniquely by its action on the basis set $\mathbf{x}_1, \ldots, \mathbf{x}_{n+2}$ we see that $g = g'h$. Let h be represented by a matrix H, which may be chosen with the correct sign such that $G = G'H$. We define constants w_i such that $H\hat{\mathbf{x}}_i = w_i \hat{\mathbf{x}}'_i$. It follows that $\eta_i = \eta'_i w_i$, since

$$\eta_i \hat{\mathbf{e}}_i = G\hat{\mathbf{x}}_i = G'H\hat{\mathbf{x}}_i = w_i G'\hat{\mathbf{x}}'_i = w_i \eta'_i \hat{\mathbf{e}}_i \ .$$

This situation is represented by the following commutative diagram.



Now, if H represents a quasi-affine transformation, then all w_i have the same sign by Proposition 4.6. We may write $w_i \doteq \epsilon$ from which one sees that $\eta_i \doteq \epsilon\eta'_i$ for all i, and the cheiral sequences of the points $\mathbf{x}_i$ and $\mathbf{x}'_i$ differ at most by a sign change.

Conversely, suppose that $\eta_i \doteq \epsilon\eta'_i$. Then $\epsilon \doteq \eta_i/\eta'_i$, since by hypothesis $\eta_i \neq 0$, and so $\eta'_i \neq 0$. On the other hand, from $\eta_i = w_i\eta'_i$ we deduce that $w_i = \eta_i/\eta'_i \doteq \epsilon$ and the w_i all have the same sign, as required. □

This theorem is not true without the assumption that $\eta_i \neq 0$, as the reader is left to discover. In practice, because of measurement inaccuracies, it will (virtually) never be the case that a computed value of η_i will equal exactly 0. Therefore, for readability in displaying cheiral sequences the practice will be adopted of writing 0 instead of -1, so that the cheiral sequence becomes a sequence of 0 and 1 values, and may be interpreted as a binary integer if desired.

11.2 The cheiral sequence in two dimensions

To illustrate the principle of the cheiral sequence, we illustrate it for sets of 4 points in the plane. The interpretation of the cheiral sequence in this way for 2-dimensional sets was suggested by Charles Rothwell. We assume that no three of the points are collinear. Let the points be $\mathbf{u}_1, \ldots, \mathbf{u}_4$. We define a particular line in the plane as follows. Denote the line through two points $\mathbf{u}_i$ and $\mathbf{u}_j$ by $< \mathbf{u}_i, \mathbf{u}_j >$. Furthermore, denote the intersection of two lines by the symbol $\times$. Thus $< \mathbf{u}_1, \mathbf{u}_2 > \times < \mathbf{u}_3, \mathbf{u}_4 >$ is the intersection of the line through $\mathbf{u}_1$ and $\mathbf{u}_2$ with the line through the points $\mathbf{u}_3$ and $\mathbf{u}_4$.

Now, construct the points $\mathbf{p}_{1234} =< \mathbf{u}_1, \mathbf{u}_2 > \times < \mathbf{u}_3, \mathbf{u}_4 >$ and $\mathbf{p}_{1324} =< \mathbf{u}_1, \mathbf{u}_3 > \times < \mathbf{u}_2, \mathbf{u}_4 >$. Then construct the line $\boldsymbol{\pi} =< \mathbf{p}_{1234}, \mathbf{p}_{1324} >$ joining these two points. This construction is shown in Fig 11 for several configurations of four points.

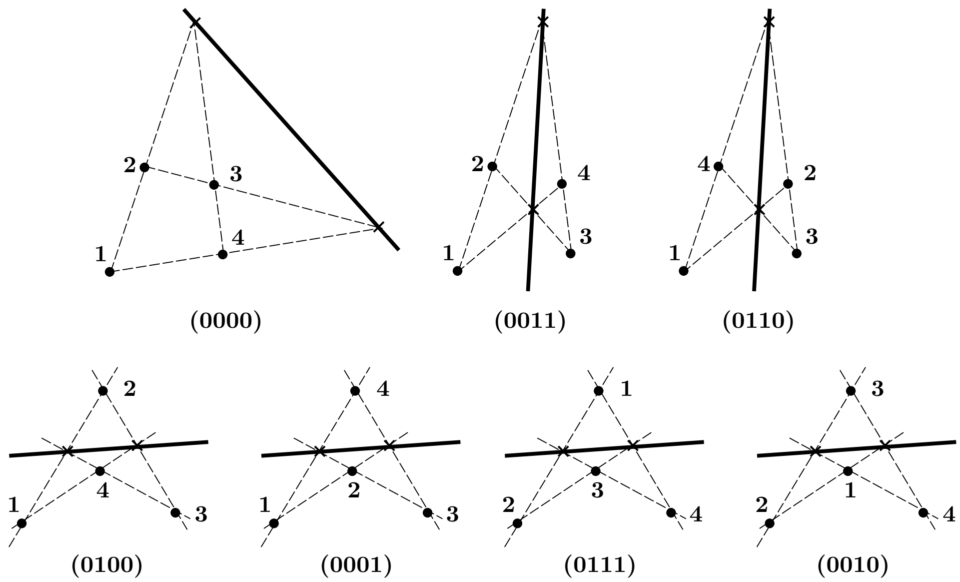


Figure 11: **Cheiral sequence in two dimensions.** *The cheiral sequnece is the sequence ξ_i for $i = 1, \ldots, 4$ where ξ_i is 0 or 1 according to whether the point $\mathbf{u}_i$ lies on the same side or the opposite side of $\boldsymbol{\pi}_\infty$ from $\mathbf{u}_1$. Shown are the 7 distinct arrangements of 4 points in the plane.*

If points $\mathbf{u}_i$ are the points of a canonical basis with homogeneous coordinates $(0,0,1)$, $(1,0,1)$, $(0,1,1)$ and $(1,1,1)$, then points $\mathbf{p}_{1234}$ and $\mathbf{p}_{1324}$ are two points on the line at infinity, and so the line $\boldsymbol{\pi}$ is the line at infinity, denoted L_∞. If on the other hand, the points $\mathbf{u}_i$ are not the points of this canonical basis, but are mapped to that basis by a projective transformation h, then the line $\boldsymbol{\pi}$ is mapped to the line at infinity. Thus, we have $\boldsymbol{\pi} = \boldsymbol{\pi}_\infty = h^{-1}(L_\infty)$, and so $\boldsymbol{\pi}$ is the line defined in the definition of the cheiral sequence. If we choose ξ_i to be ± 1 according to which side of $\boldsymbol{\pi}$ the point $\mathbf{u}_i$ lies. The sequence of values ξ_i is the cheiral sequence. It is invariant up to simultaneous reversal of all signs. The invariant values are shown in Fig 11, where for readability the digit 0 is used instead of -1. The values of ξ_i are normalized in all cases so that $\xi_1 = 0$.

As seen in the diagram (and proven by Theorem 11.25) the cheiral sequence distinguishes all non-equivalent configurations of four points. These seven configurations of points in the plane were also considered by Morin (/citemorin93a,morin94a) who found them very useful for helping distinguish point sets in the plane using projective invariants. In that work it was shown that considering the quasi-affine structure (using the present terminology) of the set of points significantly increased the capability of distinguishing point sets in the plane as compared with using only projective geometric techniques.

11.3 Computation of 3D invariants

Computation of the cheiral sequence of a set of points seen in a set of views is relatively straightforward. It takes place in four steps

1. Compute a projective reconstruction of the point set from the images.
2. Transform the projective reconstruction to a quasi-affine reconstruction.
3. Determine the mapping that maps the first five points to the canonical basis $\mathbf{e}_i$.
4. Project each point and compute the coefficients η_i.

Many ways ([1, 3, 8]) have been given for carrying out the first step of projective reconstruction. It will be easiest if one uses a method (for example [3]) that results in one of the cameras having matrix $(I \mid 0)$. Then one carries out the second step of quasi-affine reconstruction simply by swapping the last two coordinates of each point. Otherwise, the method of section 4 is still fairly straight-forward.

One may ask how many quasi-affinely distinct configurations of five points in space exist, analogous to the seven configurations of four points in the plane. We ignore configurations in which four points lie in a plane. In this case, the cheiral sequence of five points is of length five. Up to a common sign change, there are therefore 16 distinct cheiral sequences for five points. This gives an upper bound on the number of distinct configurations.

One may get an exact count by enumerating the different possible geometries of the convex hull of the points. As in two dimensions, there are two different types of configuration – those in which all five points lie on the convex hull, and those in which only 4 points lie on the convex hull. In this second case the convex hull is a tetrahedron containing the fifth point in the interior. Corresponding to the five possible choices of which point is in the interior, there are five possible such configurations.

We now analyze the configurations in which all five points lie on the convex hull. The convex hull is a polyhedron, bounded by triangular faces, since no four points are coplanar. Let n be the number of faces. Since each face has three edges, and each edge belongs to two faces, we see that there are $3n/2$ edges, and so the Euler characteristic of the polyhedron is $5 - 3n/2 + n = 2$, since the boundary of the convex hull is topologically a sphere. From this it follows that there are $n = 6$ faces and 9 edges. Since each edge meets two vertices, the sum of degrees of the vertices must equal 18. Since no vertex can have degree 5 (there are only five vertices in total), the only possibility is that there are three vertices with degree 4 and two vertices with degree 3. The polyhedron must have the shape of two tetrahedra joined along one face. There are 10 possible such configurations corresponding to the 10 different ways of choosing the two vertices with degree 3.

In total therefore there are $15 = 5 + 10$ quasi-affinely distinct configurations of five (numbered) points in three dimensions. Proposition 11.25 shows that these configurations may be distinguished by their cheiral sequences. Curiously enough, 15 is one less than the upper bound of 16 distinct cheiral sequences. Just as in the two dimensional case, there is one cheiral sequence which can not occur. Does this observation hold in higher dimensions also ? This question is left for the interested reader to resolve.

12 Experimental results

In considering real images of 3-D configurations it is necessary to take into account the effects of noise. In some cases, a value of η_i used in computing the cheiral sequence will lie so close to 0 variations due to noise can swap its sign. For robust evaluation of a cheiral sequence value, it is necessary to select a noise model and determine how errors in the input data affect the sign of each η_i. In the following discussion, noise effects are ignored, however. As usual, cheiral sequences are written using the digit 0 instead of 1, for readability.

In [4] projective invariants of 3D point sets were discussed. As an experiment in that paper, a set of images of some model houses were acquired. Fig 12 shows the three images as well as certain numbered vertices selected by hand from among those detected automatically.

Six sets of six points were chosen as in the following table which shows the indices of the points as given in Fig 12.

$$\begin{array}{lcl} S_1 &=& \{1, 2, 3, 6, 9, 10\} \;, \\ S_2 &=& \{2, 4, 6, 8, 10, 12\} \;, \\ S_3 &=& \{1, 3, 5, 7, 9, 11\} \;, \\ S_4 &=& \{1, 2, 3, 6, 7, 8\} \;, \\ S_5 &=& \{1, 4, 7, 10, 13, 12\} \;, \\ S_6 &=& \{2, 5, 8, 11, 12, 13\} \end{array}$$

From image correspondences in two views (the left two images of Fig 12) the fundamental matrix F was found and a weak realization $(P, P', \{\mathbf{x}_i\})$ was computed. For each of the six sets of indices i

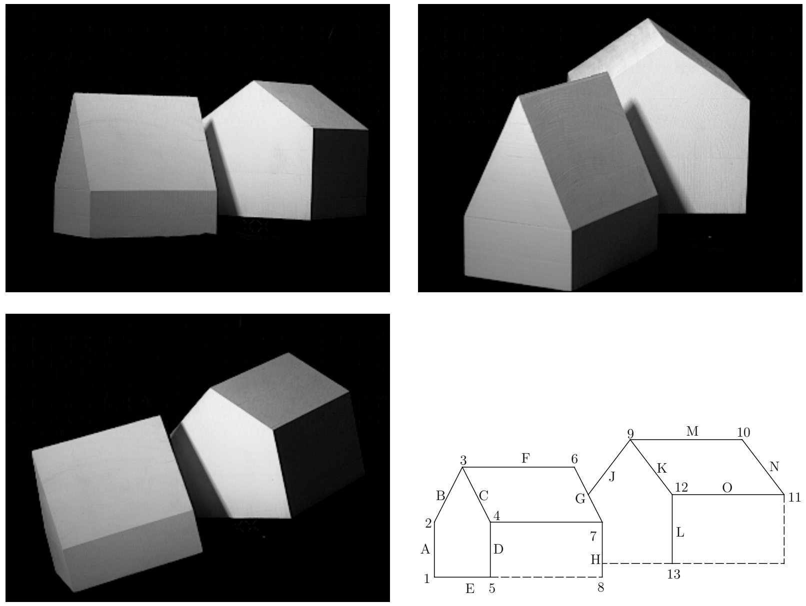


Figure 12: *Three views of houses, and numbered selected vertices*

shown above a complete projective invariant of the points $\{x_i\}$ was computed by mapping the first five points onto a canonical basis. The coordinates of the mapped sixth point constitute a projective invariant of the set of six points.

This computation was repeated with a different pair of views (right two images of Fig 12). Theory predicts that the invariants should have the same value when computed from different views, and should distinguish between non-equivalent point sets.

Table (6) shows the comparison of the computed invariant values.

0.026	0.970	0.975	0.619	0.847	0.823
0.995	**0.015**	0.064	0.841	0.252	0.548
0.967	0.066	**0.013**	0.863	0.276	0.516
0.617	0.830	0.873	**0.016**	0.704	0.752
0.861	0.238	0.289	0.708	**0.005**	0.590
0.828	0.544	0.519	0.719	0.574	**0.026**

(6)

The (i, j)-th entry of the table shows the distance according to an appropriate metric between the invariant of set S_i as computed from the first image pair with that of set S_j as computed from the second image pair. The diagonal entries of the matrix (in bold) should be close to 0.0, which indicates that the invariants had the same value when computed from different pairs of views.

Although the projective invariants computed here are quite effective at discriminating between different point sets, indicated by the fact that most off-diagonal entries are not close to zero, entries $(2, 3)$ and $(3, 2)$ are small indicating that the point sets numbered 2 and 3 are close to being equivalent up to projectivity.

Next, the cheiral sequence for each of the point sets were computed from the weak realization using the method described here. The computed values for each of the six point sets were as follows. The binary integer interpretation of the cheiral sequence is given in brackets.

$$\begin{aligned}
\chi(S_1) &= 011100 = (28)_{10} \\
\chi(S_2) &= 110000 = (60)_{10} \\
\chi(S_3) &= 000100 = (4)_{10}
\end{aligned}$$

$$\begin{array}{rcccl} \chi(S_4) & = & 111100 & = & (60)_{10} \\ \chi(S_5) & = & 101010 & = & (42)_{10} \\ \chi(S_6) & = & 100100 & = & (36)_{10} \end{array}$$

As expected these invariant values were the same whether computed using the first pair of views or the second pair. Note that the cheirality invariant clearly distinguishes point sets 2 and 3. Point sets S_2 and S_4 have the same cheiral sequence, but these are well distinguished by their projective invariants.

Conclusions : These results show that the cheiral sequence is quite effective at distinguishing between arbitrary sets of points. Given the relative ease with which the cheiral sequence may be computed, it may be extremely useful in grouping points. In addition, it may conveniently be used as an indexing function in an object recognition system. It has been demonstrated that the cheiral sequence gives supplementary information that is not available in projective invariants. As a theoretical tool, the cheiral sequence provides conditions under which image point matches may be realized by real point configurations.

References

[1] O. D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Computer Vision - ECCV '92, LNCS-Series Vol. 588, Springer-Verlag*, pages 563 – 578, 1992.

[2] R. Hartley. Invariants of points seen in multiple images. unpublished report, May 1992.

[3] R. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pages 761–764, 1992.

[4] R. I. Hartley. Invariants of lines in space. In *Proc. DARPA Image Understanding Workshop*, pages 737–744, 1993.

[5] R. I. Hartley and A. Kawauchi. Polynomials of amphicheiral knots. *Math. Ann*, 243:63 – 70, 1979.

[6] Richard I. Hartley. Euclidean reconstruction from uncalibrated views. In *Proc. of the Second Europe-US Workshop on Invariance, Ponta Delgada, Azores*, pages 187–202, October 1993.

[7] H.C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, Sept 1981.

[8] R. Mohr, F. Veillon, and L. Quan. Relative 3D reconstruction using multiple uncalibrated images. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pages 543 – 548, 1993.

[9] L. Morin. *Quelques Contributions des Invariants Projectifs à la Vision par Ordinateur*. PhD thesis, Institut National Polytechnique de Grenoble, January 1993.

[10] L. Morin, P. Brand, and R. Mohr. Indexing with projective invariants. In *Proceedings of the Syntactical and Structural Pattern Recognition workshop, Nahariya, Israel*. World Scientific Pub., 1995.

[11] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 1988.

[12] L. Robert and O.D. Faugeras. Relative 3D positioning and 3D convex hull computation from a weakly calibrated stereo pair. In *Proc. International Conference on Computer Vision*, pages 540–544, 1993.

[13] Charles A. Rothwell, Andrew Zisserman, David A. Forsyth, and Joseph L. Mundy. Canonical frames for planar object recognition. In *Computer Vision - ECCV '92, LNCS-Series Vol. 588, Springer-Verlag*, pages 757 – 772, 1992.

[14] Gunnar Sparr. Depth computations from polyhedral images. In *Computer Vision - ECCV '92, LNCS-Series Vol. 588, Springer-Verlag*, pages 378–386, 1992.

[15] I.E. Sutherland. Sketchpad: A man-machine graphical communications system. Technical Report 296, MIT Lincoln Laboratories, 1963. Also published by Garland Publishing Inc, New York, 1980.

[16] S. Wolfram. *Mathematica : A System for Doing Mathematics by Computer*. Addison-Wesley, Redwood City, California, 1988.