This chapter describes how and to what extent the spatial layout of a scene and the cameras can be recovered from two views. Suppose that a set of image correspondences **x***i ↔* **x** *i* are given. It is assumed that these correspondences come from a set of 3D points **X***i*, which are unknown. Similarly, the position, orientation and calibration of the cameras are not known. The reconstruction ask is to find the camera matrices P and P , as well as the 3D points **X***i* such that

Este capítulo describe cómo y en qué medida se puede recuperar el diseño espacial de una escena y las cámaras desde dos vistas. Supongamos que se da un conjunto de correspondencias de imágenes xi ↔ x i. Se supone que estas correspondencias provienen de un conjunto de puntos 3D Xi, que son desconocidos. Del mismo modo, la posición, orientación y calibración de las cámaras no se conocen. La solicitud de reconstrucción es encontrar las matrices de cámara P y P, así como los puntos 3D Xi de modo que

**x***i* = P**X***i* **x** *i* = P **X***i* for all *i.*

Given too few points, this task is not possible. However, if there are sufficiently many point correspondences to allow the fundamental matrix to be computed uniquely, then the scene may be reconstructed up to a projective ambiguity. This is a very significant result, and one of the major achievements of the uncalibrated approach. The ambiguity in the reconstruction may be reduced if additional information is supplied on the cameras or scene. We describe a two-stage approach where the ambiguity is first reduced to affine, and second to metric; each stage requiring information of the appropriate class.

N-VIEW

This chapter introduces the quadrifocal tensor *Qijkl* between four views, which is the analogue of the fundamental matrix for two and the trifocal tensor for three views. The quadrifocal tensor encapsulates the relationships between imaged points and lines seen in four views. It is shown that multiple view relations may be derived directly and uniformly from the intersection properties of back-projected lines and points. From this analysis the fundamental matrix F, trifocal tensor *Tijk*, and quadrifocal tensor *Qijkl* appear in a common framework involving matrix determinants. Specific formulae are given for each of these tensors in terms of the camera matrices. We also develop general counting arguments for the degrees of freedom of the tensors and the number of point and line correspondences required for tensor computation. These are given for configurations in general position and for the important special case where four or more of the elements are coplanar.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Auto-calibration is the process of determining internal camera parameters directly from multiple uncalibrated images. Once this is done, it is possible to compute a metric reconstruction from the images. Auto-calibration avoids the onerous task of calibrating cameras using special calibration objects. This gives great flexibility since, for example, a camera can be calibrated directly from an image sequence despite unknown motion and changes in some of the internal parameters. The root of the method is that a camera moves rigidly, so the absolute conic is fixed under the motion. Conversely, then, if a unique fixed conic in 3-space can be determined in some way from the images, this identifies Ω*∞*. As we have seen in earlier chapters, once Ω*∞* is identified, the metric geometry can be computed. An array of auto-calibration methods are available for this task of identifying Ω*∞*. This chapter has four main parts. First we lay out the algebraic structure of the autocalibration problem, and show how the auto-calibration equations are generated from constraints on the internal or external parameters. Second, we describe several *direct* methods for auto-calibration which involve computing the absolute conic or its image. These include estimating the absolute dual quadric over many views, or the Kruppa equations from view pairs. Third, are *stratified* methods for auto-calibration which involve two steps – first solving for the plane at infinity, then using this to solve for the absolute conic. The fourth part covers a number of special configurations including a camera rotating about its centre, a camera undergoing planar motion, and the motion of a stereo rig.

\_\_\_

Auto- (or self-) calibration is the computation of metric properties of the cameras and/or the scene from a set of uncalibrated images. This differs from conventional calibration where the camera calibration matrix K is determined from the image of a known calibration grid (chapter 7) or properties of the scene, such as vanishing points of orthogonal directions (chapter 8). Instead, in auto-calibration the metric properties are determined directly from constraints on the internal and/or external parameters. For example, suppose we have a set of images acquired by a camera with fixed internal parameters, and that a projective reconstruction is computed from point correspondences across the image set. The reconstruction computes a projective camera matrix P*i* for each view. Our constraint is that for the actual cameras the internal parameter matrix K is the same (but unknown) for each view. Now, each camera P*i* of the projective reconstruction may be decomposed as P*i* = K*i*[R*i |* **t***i*] but in general the calibration matrix K*i* will differ for each view. Thus the constraint will *not* be satisfied by the projective reconstruction. However, we have the freedom to vary our projective reconstruction by transforming the camera matrices by a homography H. Since the actual cameras have fixed internal parameters, there will exist a homography (or a family of homographies) such that the transformed cameras P*i*H do decompose as P*i*H = KR*i*[I *|* **t***i*], with the same calibration matrix for each camera, so the reconstruction is consistent with the constraint. Provided there are sufficiently many views and the motion between the views is general (see later), then this consistency constrains H to the extent that the reconstruction transformed by H is within a similarity transformation of the actual cameras and scene, i.e. we achieve a metric reconstruction.