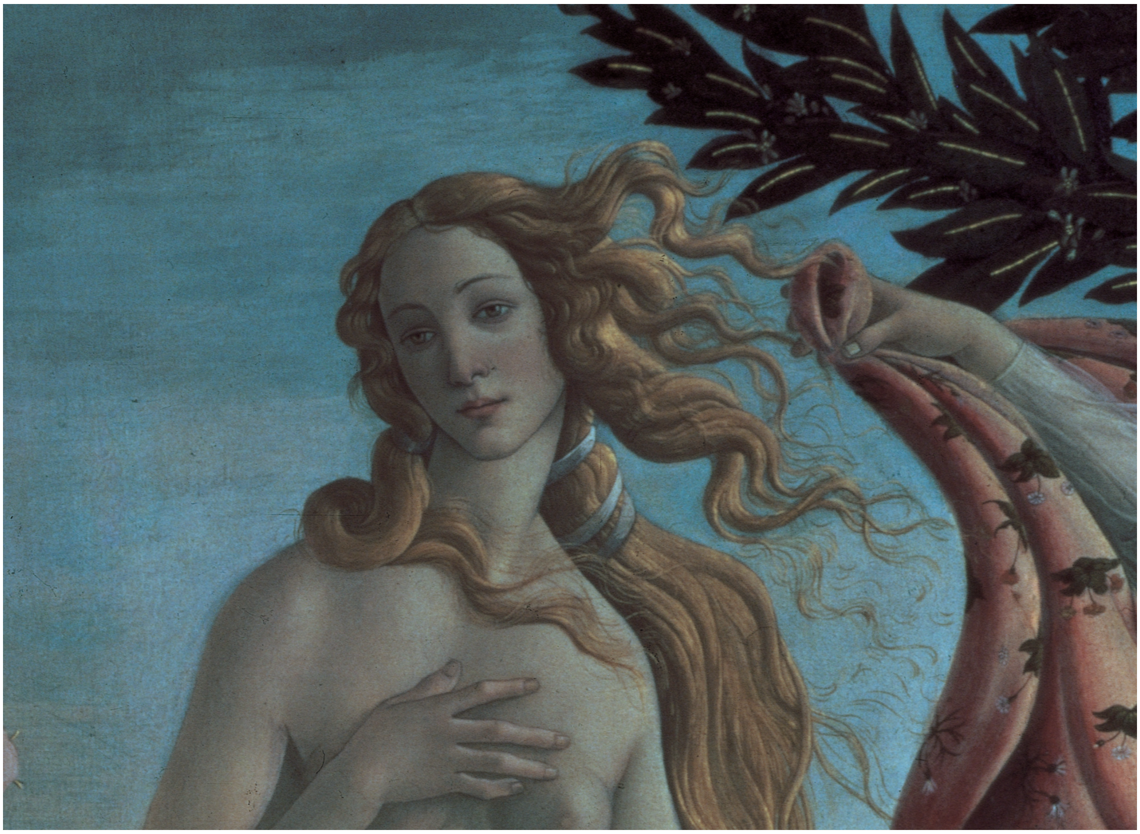
Part II

Two-View Geometry



The Birth of Venus (detail), c. 1485 (tempera on canvas) by Sandro Botticelli (1444/5-1510)
Galleria degli Uffizi, Florence, Italy/Bridgeman Art Library

Outline

This part of the book covers the geometry of two perspective views. These views may be acquired simultaneously as in a stereo rig, or acquired sequentially, for example by a camera moving relative to the scene. These two situations are geometrically equivalent and will not be differentiated here. Each view has an associated camera matrix, $\mathtt{P}, \mathtt{P}'$, where $'$ indicates entities associated with the second view, and a 3-space point $\mathbf{X}$ is imaged as $\mathbf{x} = \mathtt{P}\mathbf{X}$ in the first view, and $\mathbf{x}' = \mathtt{P}'\mathbf{X}$ in the second. Image points $\mathbf{x}$ and $\mathbf{x}'$ *correspond* because they are the image of the same 3-space point. There are three questions that will be addressed:

(i) **Correspondence geometry.** Given an image point $\mathbf{x}$ in the first view, how does this constrain the position of the corresponding point $\mathbf{x}'$ in the second view?

(ii) **Camera geometry (motion).** Given a set of corresponding image points $\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i\}, i = 1, \ldots, n$, what are the cameras $\mathtt{P}$ and $\mathtt{P}'$ for the two views?

(iii) **Scene geometry (structure).** Given corresponding image points $\mathbf{x} \leftrightarrow \mathbf{x}'$ and cameras $\mathtt{P}, \mathtt{P}'$, what is the position of (their pre-image) $\mathbf{X}$ in 3-space?

Chapter 9 describes the *epipolar geometry* of two views, and directly answers the first question: a point in one view defines an epipolar line in the other view on which the corresponding point lies. The epipolar geometry depends only on the cameras – their relative position and their internal parameters. It does *not* depend at all on the scene structure. The epipolar geometry is represented by a 3×3 matrix called the *fundamental matrix* $\mathtt{F}$. The anatomy of the fundamental matrix is described, and its computation from camera matrices $\mathtt{P}$ and $\mathtt{P}'$ given. It is then shown that $\mathtt{P}$ and $\mathtt{P}'$ may be computed from $\mathtt{F}$ up to a projective ambiguity of 3-space.

Chapter 10 describes one of the most important results in uncalibrated multiple view geometry – a *reconstruction* of both cameras and scene structure can be computed from image point correspondences alone; no other information is required. This answers both the second and third questions simultaneously. The reconstruction obtained from point correspondences alone is up to a projective ambiguity of 3-space, and this ambiguity can be resolved by supplying well defined additional information on the cameras or scene. In this manner an affine or metric reconstruction may be computed from uncalibrated images. The following two chapters then fill in the details and numerical algorithms for computing this reconstruction.

Chapter 11 describes methods for computing $\mathtt{F}$ from a set of corresponding image points $\{\mathbf{x}_i \leftrightarrow \mathbf{x}'_i\}$, even though the structure (3D pre-image $\mathbf{X}_i$) of these points is unknown and the camera matrices are unknown. The cameras $\mathtt{P}$ and $\mathtt{P}'$ may then be determined, up to a projective ambiguity, from the computed $\mathtt{F}$.

Chapter 12 then describes the computation of scene structure by *triangulation* given the cameras and corresponding image points – the point $\mathbf{X}$ in 3-space is computed as the intersection of rays back-projected from the corresponding points $\mathbf{x}$ and $\mathbf{x}'$ via their associated cameras $\mathtt{P}, \mathtt{P}'$. Similarly, the 3D position of other geometric entities, such as lines or conics, may also be computed given their image correspondences.

Chapter 13 covers the two-view geometry of planes. It provides an alternative answer to the first question: if scene points lie on a plane, then once the geometry of this plane is computed, the image $\mathbf{x}$ of a point in one image determines the position of $\mathbf{x}'$ in the other image. The points are related by a plane projective transformation. This chapter also describes a particularly important projective transformation between views – the *infinite homography*, which is the transformation arising from the plane at infinity.

Chapter 14 describes two-view geometry in the specialized case that the two cameras $\mathtt{P}$ and $\mathtt{P}'$ are affine. This case has a number of simplifications over the general projective case, and provides a very good approximation in many practical situations.