7 Image Formation

7.1 Introduction

Image formation includes three major aspects. One is *geometric* in nature. The question is *where* we find an object in the image. Essentially, all imaging techniques project a three-dimensional space in one way or the other onto a two-dimensional image plane. Thus, basically, imaging can be regarded as a projection from 3-D into 2-D space. The loss of one coordinate constitutes a severe loss of information about the geometry of the observed scene. However, we unconsciously and constantly experience that our visual system perceives a three-dimensional impression sufficiently well that we can grasp the three-dimensional world around us and interact with it. The ease with which this reconstruction task is performed by biological visual systems might tempt us to think that this is a simple task. But — as we will see in Chapters 8 and 17 — it is not that simple.

The second aspect is *radiometric* in nature. How "bright" is an imaged object, and how does the brightness in the image depend on the optical properties of the object and the image formation system? The radiometry of an imaging system is discussed in Section 7.5. For the basics of radiometry see Section 6.2.

The third question is, finally, what happens to an image when we represent it with an array of digital numbers to process it with a digital computer? How do the processes that transform a continuous image into such an array — known as *digitization* and *quantization* — limit the resolution in the image or introduce artifacts? These questions are addressed in Chapter 9.

7.2 World and Camera Coordinates

7.2.1 Definition

Basically, the position of objects in 3-D space can be described in two different ways (Fig. 7.1). First, we can use a coordinate system that is related to the scene observed. These coordinates are called *world coordinates* and denoted as $\boldsymbol{X}' = [X_1', X_2', X_3']^T$. The X_1' and X_2' coordinates describe the horizontal and X_3' the vertical positions, respectively. Sometimes, an

B. Jähne, Digital Image Processing
ISBN 3-540-24035-7

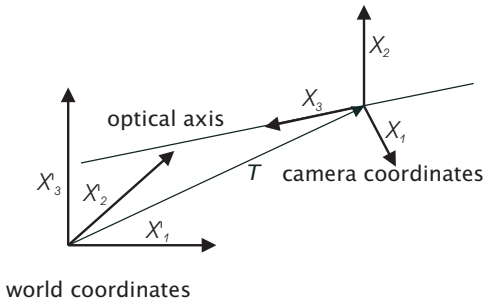

Figure 7.1: *Illustration of world and camera coordinates.*

alternative convention with non-indexed coordinates $\boldsymbol{X}' = [X', Y', Z']^T$ is more convenient. Both notations are used in this book.

A second system, the *camera coordinates* $\boldsymbol{X} = [X_1, X_2, X_3]^T$, can be fixed to the camera observing the scene. The X_3 axis is aligned with the *optical axis* of the camera system (Fig. 7.1). Physicists are familiar with such considerations. It is common to discuss physical phenomena in different coordinate systems. In elementary mechanics, for example, motion is studied with respect to two observers, one at rest, the other moving with the object.

Transition from world to camera coordinates generally requires a *translation* and a *rotation*. First, we shift the origin of the world coordinate system to the origin of the camera coordinate system by the translation vector $\boldsymbol{T}$ (Fig. 7.1). Then we change the orientation of the shifted system by rotations about suitable axes so that it coincides with the camera coordinate system. Mathematically, translation can be described by vector subtraction and rotation by multiplication of the coordinate vector with a matrix:

$$\boldsymbol{X} = \boldsymbol{R}(\boldsymbol{X}' - \boldsymbol{T}). \tag{7.1}$$

7.2.2 Rotation

Rotation of a coordinate system has two important features. It does not change the length or *norm* of a vector and it keeps the coordinate system orthogonal. A transformation with these features is known in linear algebra as an *orthonormal* transform.

The coefficients in a transformation matrix have an intuitive meaning. This can be seen when we apply the transformation to unit vectors $\bar{\boldsymbol{E}}_p$ in the direction of the coordinate axes. With $\bar{\boldsymbol{E}}_1$, for instance, we obtain

$$\bar{\boldsymbol{E}}_1' = \boldsymbol{A}\bar{\boldsymbol{E}}_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix}. \tag{7.2}$$

Thus, the columns of the transformation matrix give the coordinates of the base vectors in the new coordinate system. Knowing this property, it is easy to formulate the orthonormality condition that has to be met by the rotation matrix R:

$$\boldsymbol{R}^T\boldsymbol{R} = \boldsymbol{I} \quad \text{or} \quad \sum_{m=1}^{3} r_{km}r_{lm} = \delta_{k-l}, \tag{7.3}$$

where $\boldsymbol{I}$ denotes the identity matrix, whose elements are one and zero on diagonal and non-diagonal positions, respectively. Using Eq. (7.2), this equation simply states that the transformed base vectors remain orthogonal:

$$\bar{\boldsymbol{E}}_k'^T\bar{\boldsymbol{E}}_l' = \delta_{k-l}. \tag{7.4}$$

Equation Eq. (7.3) leaves three matrix elements independent out of nine. Unfortunately, the relationship between the matrix elements and three parameters to describe rotation turns out to be quite complex and nonlinear. A common procedure involves the three Eulerian rotation angles (ϕ, θ, ψ). A lot of confusion exists in the literature about the definition of the Eulerian angle. We follow the standard mathematical approach. We use right-hand coordinate systems and count rotation angles positive in the counterclockwise direction. The rotation from the shifted world coordinate system to the camera coordinate system is decomposed into three steps (see Fig. 7.2, [60]).

1. Rotation about X_3' axis by the angle ϕ, $\boldsymbol{X}'' = \boldsymbol{R}_\phi\boldsymbol{X}'$:

$$\boldsymbol{R}_\phi = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7.5}$$

2. Rotation about X_1'' axis by θ, $\boldsymbol{X}''' = \boldsymbol{R}_\theta\boldsymbol{X}''$:

$$\boldsymbol{R}_\theta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \tag{7.6}$$

3. Rotation about X_3''' axis by ψ, $\boldsymbol{X} = \boldsymbol{R}_\psi\boldsymbol{X}'''$:

$$\boldsymbol{R}_\psi = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7.7}$$

Cascading the three rotations, $\boldsymbol{R}_\psi\boldsymbol{R}_\theta\boldsymbol{R}_\phi$, yields the matrix

$$\begin{bmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & \sin\theta\sin\psi \\ -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & \sin\theta\cos\psi \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{bmatrix}.$$

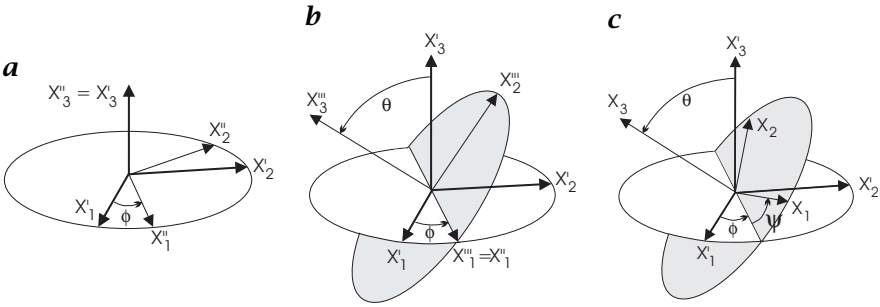

Figure 7.2: *Rotation of world coordinates $\boldsymbol{X}'$ to camera coordinates $\boldsymbol{X}$ using the three Eulerian angles (ϕ, θ, ψ) with successive rotations about the* ***a*** X_3', ***b*** X_1'', *and* ***c*** X_3''' *axes.*

The inverse transformation from camera coordinates to world coordinates is given by the transpose of the above matrix. Since matrix multiplication is not commutative, rotation is also not commutative. Therefore, it is important not to interchange the order in which rotations are performed.

Rotation is only commutative in the limit of an infinitesimal rotation. Then, the cosine and sine terms reduce to 1 and ε, respectively. This limit has some practical applications since minor rotational misalignments are common.

Rotation about the X_3 axis, for instance, can be

$$\boldsymbol{X} = \boldsymbol{R}_\varepsilon \boldsymbol{X}' = \begin{bmatrix} 1 & \varepsilon & 0 \\ -\varepsilon & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{X}' \quad \text{or} \quad \begin{aligned} X_1 &= X_1' + \varepsilon X_2' \\ X_2 &= X_2' - \varepsilon X_1' \\ X_3 &= X_3' \end{aligned} .$$

As an example we discuss the rotation of the point $[X_1', 0, 0]^T$. It is rotated into the point $[X_1', \varepsilon X_1', 0]^T$ while the correct position would be $[X_1' \cos\varepsilon, X_1' \sin\varepsilon, 0]^T$. Expanding the trigonometric function in a Taylor series to third order yields a position error of $[1/2\varepsilon^2 X_1', 1/6\varepsilon^3 X_1', 0]^T$. For a 512×512 image ($X_1' < 256$ for centered rotation) and an error limit of less than 1/20 pixel, ε must be smaller than 0.02 or 1.15°. This is still a significant rotation vertically displacing rows by up to $\pm\varepsilon X' = \pm 5$ pixels.

7.3 Ideal Imaging: Perspective Projection

7.3.1 The Pinhole Camera

The basic geometric aspects of image formation by an optical system are well modeled by a *pinhole camera*. The imaging element of this camera

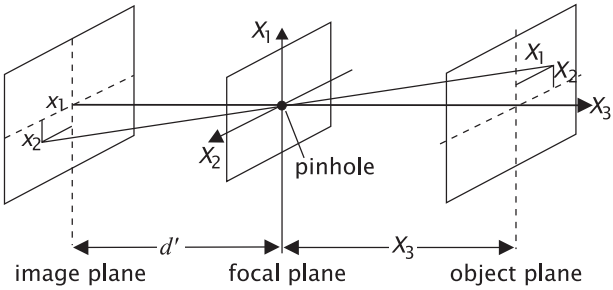

Figure 7.3: *Image formation with a pinhole camera.*

is an infinitesimally small hole (Fig. 7.3). The single light ray coming from a point of the object at $[X_1, X_2, X_3]^T$ which passes through this hole meets the image plane at $[x_1, x_2, -d_i]^T$. Through this condition an image of the object is formed on the image plane. The relationship between the 3-D world and the 2-D *image coordinates* $[x_1, x_2]^T$ is given by

$$x_1 = -\frac{d' X_1}{X_3}, \quad x_2 = -\frac{d' X_2}{X_3}. \tag{7.8}$$

The two world coordinates parallel to the image plane are scaled by the factor d'/X_3. Therefore, the image coordinates $[x_1, x_2]^T$ contain only ratios of world coordinates, from which neither the distance nor the true size of an object can be inferred.

A straight line in the world space is projected onto a straight line at the image plane. This important feature can be proved by a simple geometric consideration. All light rays emitted from a straight line pass through the pinhole. Consequently they all lie on a plane that is spanned by the straight line and the pinhole. This plane intersects with the image plane in a straight line.

All object points on a ray through the pinhole are projected onto a single point in the image plane. In a scene with several transparent objects, the objects are projected onto each other. Then we cannot infer the three-dimensional structure of the scene at all. We may not even be able to recognize the shape of individual objects. This example demonstrates how much information is lost by projection of a 3-D scene onto a 2-D image plane.

Most natural scenes, however, contain opaque objects. Here the observed 3-D space is essentially reduced to 2-D surfaces. These surfaces can be described by two two-dimensional functions $g(x_1, x_2)$ and $X_3(x_1, x_2)$ instead of the general description of a 3-D scalar gray value image $g(X_1, X_2, X_3)$. A surface in space is completely projected onto the image plane provided that not more than one point of the surface lies on the same ray through the pinhole. If this condition is not met, parts

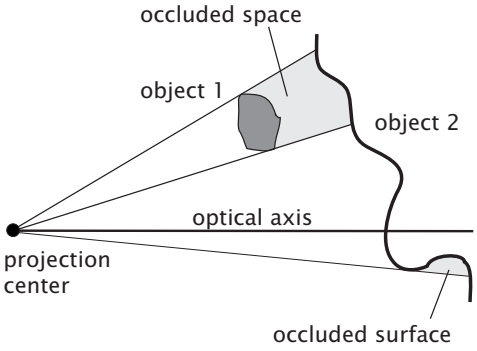

Figure 7.4: *Occlusion of more distant objects and surfaces by perspective projection.*

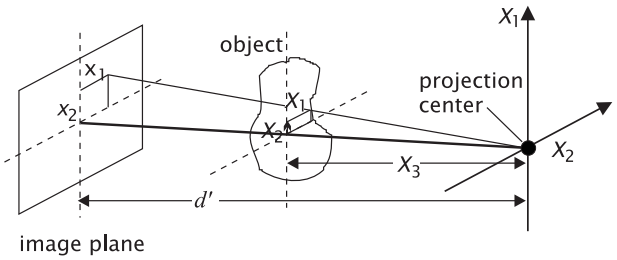

Figure 7.5: *Perspective projection with x-rays.*

of the surface remain invisible. This effect is called *occlusion*. The occluded 3-D space can be made visible if we put a point light source at the position of the pinhole (Fig. 7.4). Then the invisible parts of the scene lie in the shadow of those objects that are closer to the camera.

As long as we can exclude occlusion, we only need the depth map $X_3(x_1, x_2)$ to reconstruct the 3-D shape of a scene completely. One way to produce it — which is also used by our visual system — is by stereo imaging, i. e., observation of the scene with two sensors from different points of view (Section 8.2.1).

7.3.2 Projective Imaging

Imaging with a pinhole camera is essentially a *perspective projection*, because all rays must pass through one central point, the pinhole. Thus the pinhole camera model is very similar to imaging with penetrating rays, such as x-rays, emitted from a point source (Fig. 7.5). In this case, the object lies between the central point and the image plane.

The projection equation corresponds to Eq. (7.8) except for the sign:

$$\boxed{\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \longmapsto \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \dfrac{d' X_1}{X_3} \\ \dfrac{d' X_2}{X_3} \end{bmatrix}}. \tag{7.9}$$

The image coordinates divided by the image distance d_i are called *generalized image coordinates*:

$$\tilde{x}_1 = \frac{x_1}{d'}, \quad \tilde{x}_2 = \frac{x_2}{d'}. \tag{7.10}$$

Generalized image coordinates are dimensionless and denoted by a tilde. They are equal to the tangent of the angle with respect to the optical axis of the system with which the object is observed. These coordinates explicitly take the limitations of the projection onto the image plane into account. From these coordinates, we cannot infer absolute positions but know only the angle at which the object is projected onto the image plane. The same coordinates are used in astronomy. The general projection equation of perspective projection Eq. (7.9) then reduces to

$$\boldsymbol{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \longmapsto \tilde{\boldsymbol{x}} = \begin{bmatrix} \dfrac{X_1}{X_3} \\ \dfrac{X_2}{X_3} \end{bmatrix}. \tag{7.11}$$

We will use this simplified projection equation in all further considerations. For optical imaging, we just have to include a minus sign or, if speaking geometrically, reflect the image at the origin of the coordinate system.

7.4 Real Imaging

7.4.1 Basic Geometry of an Optical System

The model of a pinhole camera is an oversimplification of an imaging system. A pinhole camera forms an image of an object at *any* distance while a real optical system forms a sharp image only within a certain distance range. Fortunately, the geometry even for complex optical systems can still be modeled with a small modification of perspective projection as illustrated in Figs. 7.6 and 7.7. The focal plane has to be replaced by two *principal planes*. The two principal planes meet the *optical axis* at the *principal points*. A ray directed towards the first principal point appears — after passing through the system — to originate from the second principal point without angular deviation (Fig. 7.6). The distance

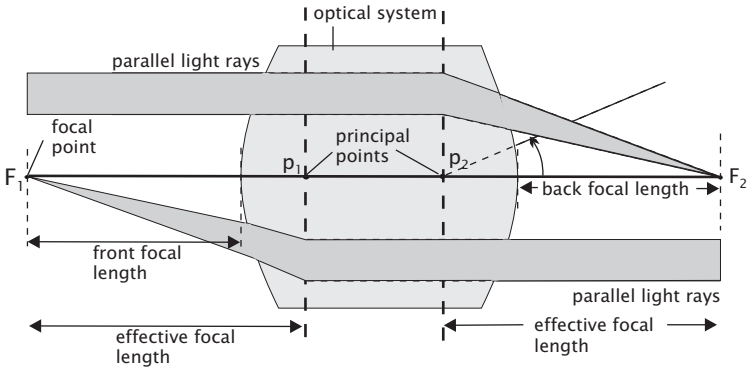

Figure 7.6: *Black box model of an optical system.*

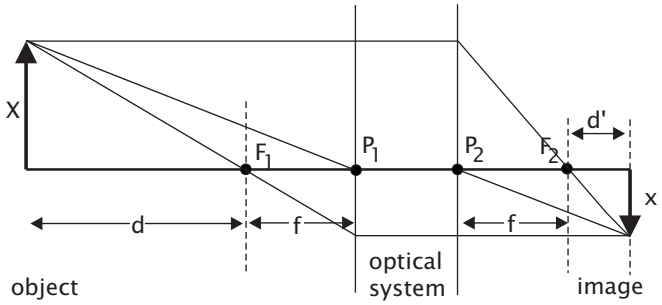

Figure 7.7: *Optical imaging with an optical system modeled by its principal points P_1 and P_2 and focal points F_1 and F_2. The system forms an image that is a distance d' behind F_2 from an object that is the distance d in front of F_1.*

between the two principal planes thus models the axial extension of the optical system.

As illustrated in Fig. 7.6, rays between the two principal planes are always parallel and parallel rays entering the optical system from left and right meet at the second and first focal point, respectively. For practical purposes, the following definitions also are useful: The *effective focal length* is the distance from the principal point to the corresponding focal point. The *front focal length* and *back focal length* are the distances from the first and last surface of the optical system to the first and second focal point, respectively.

The relation between the object distance and the image distance becomes very simple if they are measured from the focal points (Fig. 7.7),

$$\boxed{dd' = f^2.} \tag{7.12}$$

This is the Newtonian form of the *image equation*. The possibly better known Gaussian form uses the distances as to the principal points:

$$\frac{1}{d'+f}+\frac{1}{d+f}=\frac{1}{f} \tag{7.13}$$

7.4.2 Lateral and Axial Magnification

The *lateral magnification* m_l of an optical system is given by the ratio of the image size, x, to the object size, X:

$$\boxed{m_l=\frac{x}{X}=\frac{f}{d}=\frac{d'}{f}=\frac{f+d'}{f+d}.} \tag{7.14}$$

The lateral magnification m_l is proportional to d': $d'=fm_l$ and inversely proportional to d: $d'=f/m_l$. Therefore it is easy to compute the distance to the object (d) and the distance of the image plane to the focal plane (d') from a given magnification. Three illustrative examples follow: object at infinity ($m_l=0$): $d'=0$, magnification 1/10 ($m_l=1/10$): $d'=f/10$, one-to-one imaging: ($m_l=1$): $d'=d=f$.

A less well-known quantity is the *axial magnification* that relates the positions of the image plane and object plane to each other. Thus the axial magnification gives the magnification along the optical axis. If we shift a point in the object space along the optical axis, how large is the shift of the image plane? In contrast to the lateral magnification, the axial magnification is not constant with the position along the optical axis. Therefore the axial magnification is only defined in the limit of small changes. We use slightly modified object and image positions $d+\Delta X_3$ and $d'-\Delta x_3$ and introduce them into Eq. (7.12). Then a first-order Taylor expansion in ΔX_3 and Δx_3 (assuming that $\Delta X_3 \ll d$ and $\Delta x_3 \ll d'$) yields

$$\frac{\Delta x_3}{\Delta X_3}\approx\frac{d'}{d} \tag{7.15}$$

and the axial magnification m_a is given by

$$\boxed{m_a\approx\frac{d'}{d}=\frac{f^2}{d^2}=\frac{d'^2}{f^2}=m_l^2.} \tag{7.16}$$

7.4.3 Depth of Focus and Depth of Field

The image equations Eqs. (7.12) and (7.13) determine the relation between object and image distances. If the image plane is slightly shifted or the object is closer to the lens system, the image is not rendered useless. It rather gets blurred. The degree of blurring depends on the deviation from the distances given by the image equation.

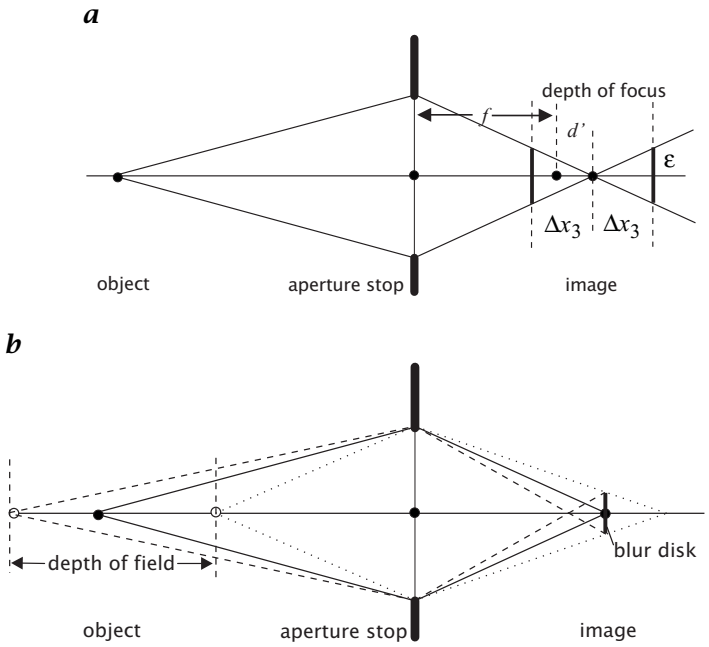

Figure 7.8: *Illustration of the* ***a*** *depth of focus and* ***b*** *depth of field with an on-axis point object.*

The concepts of depth of focus and depth of field are based on the fact that a certain degree of blurring does not affect the image quality. For digital images it is naturally given by the size of the sensor elements. It makes no sense to resolve smaller structures. We compute the blurring in the framework of geometrical optics using the image of a point object as illustrated in Fig. 7.8a. At the image plane, the point object is imaged to a point. It smears to a disk with the radius ϵ with increasing distance from the image plane. Introducing the f-number n_f of an optical system as the ratio of the focal length and diameter of lens aperture $2r$

$$n_f = \frac{f}{2r}, \tag{7.17}$$

we can express the radius of the blur disk as:

$$\epsilon = \frac{1}{2n_f}\frac{f}{f+d'}\Delta x_3, \tag{7.18}$$

where Δx_3 is the distance from the (focused) image plane. The range of positions of the image plane, $[d' - \Delta x_3, d' + \Delta x_3]$, for which the radius of the blur disk is lower than ϵ, is known as the *depth of focus*.

Equation (7.18) can be solved for Δx_3 and yields

$$\Delta x_3 = 2n_f \left(1 + \frac{d'}{f}\right) \epsilon = 2n_f(1 + m_l)\epsilon, \tag{7.19}$$

where m_l is the lateral magnification as defined by Eq. (7.14). Equation (7.19) illustrates the critical role of the n_f-number and magnification for the depth of focus. Only these two parameters determine for a given ϵ the depth of focus and depth of field.

Of even more importance for practical usage than the depth of focus is the *depth of field*. The depth of field is the range of object positions for which the radius of the blur disk remains below a threshold ϵ at a fixed image plane (Fig. 7.8b). With Eqs. (7.12) and (7.19) we obtain

$$d \pm \Delta X_3 = \frac{f^2}{d' \mp \Delta x_3} = \frac{f^2}{d' \mp 2n_f(1 + m_l)\epsilon}. \tag{7.20}$$

In the limit of $\Delta X_3 \ll d$, Eq. (7.20) reduces to

$$\boxed{\Delta X_3 \approx 2n_f \cdot \frac{1 + m_l}{m_l^2}\epsilon.} \tag{7.21}$$

If the depth of field includes the infinite distance, the minimum distance for a sharp image is

$$d_{\min} = \frac{f^2}{4n_f(1 + m_l)\epsilon} \approx \frac{f^2}{4n_f\epsilon}. \tag{7.22}$$

A typical high resolution CCD camera has sensor elements, which are about $10 \times 10\,\mu$m in size. Thus we can allow for a radius of the unsharpness disc of $5\,\mu$m. Assuming a lens with an f-number of 2 and a focal length of 15 mm, according to Eq. (7.21) we have a depth of field of $\pm$ 0.2 m at an object distance of 1.5 m, and according to Eq. (7.22) the depth of field reaches from 5 m to infinity. This example illustrates that even with this small f-number and the relatively short distance, we may obtain a large depth of field.

For high magnifications as in *microscopy*, the depth of field is very small. With $m_l \gg 1$, Eq. (7.21) reduces to

$$\Delta X_3 \approx \frac{2n_f\epsilon}{m_l}. \tag{7.23}$$

With a 50-fold enlargement ($m_l = 50$) and $n_f = 1$, we obtain the extreme low depth of field of only $0.2\,\mu$m.

Generally, the whole concept of depth of field and depth of focus as discussed here is only valid in the limit of geometrical optics. It can only be used for blurring that is significantly larger than that caused by the aberrations or diffraction of the optical system (Section 7.6.3).

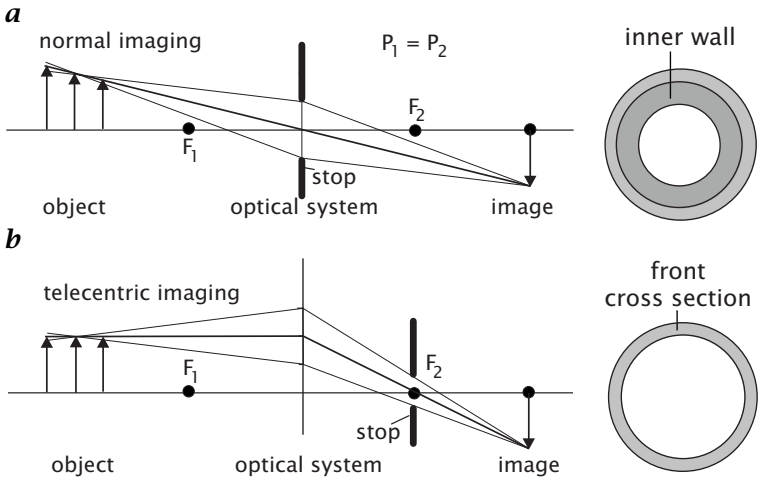

Figure 7.9: ***a*** *Standard diverging imaging with stop at the principal point;* ***b*** *telecentric imaging with stop at the second focal point. On the right side it is illustrated how a short cylindrical tube whose axis is aligned with the optical axis is imaged with the corresponding set up.*

7.4.4 Telecentric Imaging

In a standard optical system, a converging beam of light enters an optical system. This setup has a significant disadvantage for optical gauging (Fig. 7.9a). The object appears larger if it is closer to the lens and smaller if it is farther away from the lens. As the depth of the object cannot be inferred from its image, either the object must be at a precisely known depth or measurement errors are unavoidable.

A simple change in the position of the *aperture stop* from the principal point to the first focal point solves the problem and changes the imaging system to a telecentric lens (Fig. 7.9b). By placing the stop at this point, the *principal rays* (ray passing through the center of the aperture) are parallel to the optical axis in the object space. Therefore, slight changes in the position of the object do not change the size of the image of the object. The farther it is away from the focused position, the more it is blurred, of course. However, the center of the blur disk does not change the position.

Telecentric imaging has become an important principle in machine vision. Its disadvantage is, of course, that the diameter of a telecentric lens must be at least of the size of the object to be imaged. This makes telecentric imaging very expensive for large objects.

Figure 7.9 illustrates how a cylinder aligned with the optical axis with a thin wall is seen with a standard lens and a telecentric lens. Standard

imaging sees the cross-section and the inner wall and telecentric imaging the cross-section only.

The discussion of telecentric imaging emphasizes the importance of stops in the construction of optical systems, a fact that is often not adequately considered.

7.4.5 Geometric Distortion

A real optical system causes deviations from a perfect perspective projection. The most obvious *geometric distortions* can be observed with simple spherical lenses as barrel- or cushion-shaped images of squares. Even with a corrected lens system these effects are not completely suppressed.

This type of distortion can easily be understood by considerations of symmetry. As lens systems show cylindrical symmetry, concentric circles only suffer a distortion in the radius. This distortion can be approximated by

$$\boldsymbol{x}' = \frac{\boldsymbol{x}}{1 + k_3|\boldsymbol{x}|^2}. \tag{7.24}$$

Depending on whether k_3 is positive or negative, barrel- and cushion-shaped distortions in the images of squares will be observed. Commercial TV lenses show a radial deviation of several image points (pixels) at the edge of the sensor. If the distortion is corrected with Eq. (7.24), the residual error is less than 0.06 image points [119].

This high degree of correction, together with the geometric stability of modern CCD sensors, accounts for subpixel accuracy in distance and area measurements without using expensive special lenses. Lenz [120] discusses further details which influence the geometrical accuracy of CCD sensors.

Distortions also occur if non-planar surfaces are projected onto the image plane. These distortions prevail in satellite and aerial imagery. Thus correction of geometric distortion in images is a basic topic in remote sensing and photogrammetry [166].

Accurate correction of the geometrical distortions requires shifting of image points by fractions of the distance between two image points. We will deal with this problem later in Section 10.5 after we have worked out the knowledge necessary to handle it properly.

7.5 Radiometry of Imaging

If is not sufficient to know only the geometry of imaging. Equally important is to consider how the irradiance at the image plane is related to the radiance of the imaged objects and which parameters of an optical system influence this relationship. For a discussion of the fundamentals of

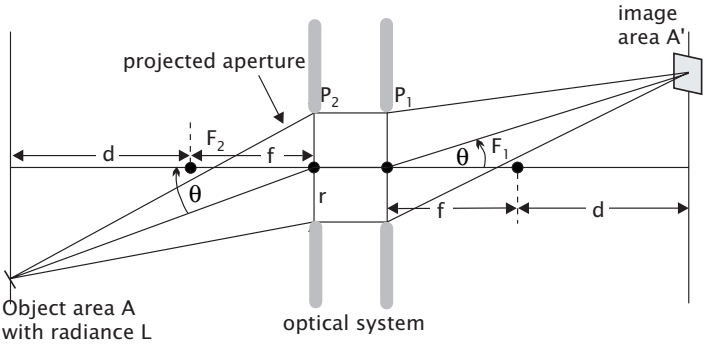

Figure 7.10: *An optical system receives a flux density that corresponds to the product of the radiance of the object and the solid angle subtended by the projected aperture as seen from the object. The flux emitted from the object area A is imaged onto the image area A'.*

radiometry, especially all terms describing the properties of radiation, we refer to Section 6.2.

The path of radiation from a light source to the image plane involves a chain of processes (see Fig. 6.1). In this section, we concentrate on the observation path (compare Fig. 6.1), i. e., how the radiation emitted from the object to be imaged is collected by the imaging system.

7.5.1 Object Radiance and Image Irradiance

An optical system collects part of the radiation emitted by an object (Fig. 7.10). We assume that the object is a homogeneous Lambertian radiator with the radiance L. The aperture of the optical system appears from the object to subtend a certain solid angle Ω. The projected circular aperture area is $\pi r^2 \cos\theta$ at a distance $(d+f)/\cos\theta$. Then, according to Eq. (6.4), a flux

$$\Phi = A\Omega L = A\frac{\pi r^2 \cos^3\theta}{(d+f)^2}L \tag{7.25}$$

enters the optical system. The radiation emitted from the area A projected onto the object plane, i. e. $A/\cos\theta$ is imaged onto the area A'. Therefore, the flux Φ must be divided by the area A' in order to compute the image irradiance E'. After Eq. (7.14), the area ratio can be expressed as

$$\frac{A/\cos\theta}{A'} = \frac{1}{m_l^2} = \frac{(f+d)^2}{(f+d')^2}. \tag{7.26}$$

We further assume that the optical system has a transmittance t. Inserting Eq. (7.26) into Eq. (7.25) finally leads to the following object ra-

diance/image irradiance relation:

$$E' = \frac{\Phi}{A'} = t\pi \left(\frac{r}{f + d'} \right)^2 \cos^4 \theta \, L. \tag{7.27}$$

This fundamental relationship states that the image irradiance is proportional to the object radiance. This is the base for the linearity of optical imaging. The optical system is described by two simple terms: its (total) transmittance t and the ratio of the aperture radius to the distance of the image from the first principal point. For distant objects $d \gg f$, $d' \ll f$, Eq. (7.27) reduces to

$$\boxed{E' = t\pi \frac{\cos^4 \theta}{4n_f^2} L, \qquad d \gg f} \tag{7.28}$$

using the f-number n_f (Eq. (7.17)). For real optical systems, equations Eqs. (7.27) and (7.28) are only an approximation. If part of the incident beam is cut off by additional apertures or limited lens diameters (*vignetting*), the fall-off is even steeper at high angles θ. On the other hand, a careful design of the position of the aperture can make the fall-off less steep than $\cos^4 \theta$. As also the residual reflectivity of the lens surfaces depends on the angle of incidence, the true fall-off depends strongly on the design of the optical system and is best determined experimentally by a suitable calibration setup.

7.5.2 Invariance of Radiance

The astonishing fact that the image irradiance is so simply related to the object radiance has its cause in a fundamental invariance. An image has a radiance just like a real object. It can be taken as a source of radiation by further optical elements. A fundamental theorem of radiometry now states that the radiance of an image is equal to the radiance of the object times the transmittance of the optical system.

The theorem can be proved using the assumption that the radiative flux Φ through an optical system is preserved except for absorption in the system leading to a transmittance less than one. The solid angles that the object and image subtend in the optical system are

$$\Omega = A_0/(d + f)^2 \quad \text{and} \quad \Omega' = A_0/(d' + f)^2, \tag{7.29}$$

where A_0 is the effective area of the aperture.

The flux emitted from an area A of the object is received by the area $A' = A(d' + f)^2/(d + f)^2$ in the image plane (Fig. 7.11a). Therefore, the radiances are

$$\begin{aligned} L &= \frac{\Phi}{\Omega A} = \frac{\Phi}{A_0 A}(d + f)^2 \\ L' &= \frac{t\Phi}{\Omega' A'} = \frac{t\Phi}{A_0 A}(d + f)^2, \end{aligned} \tag{7.30}$$

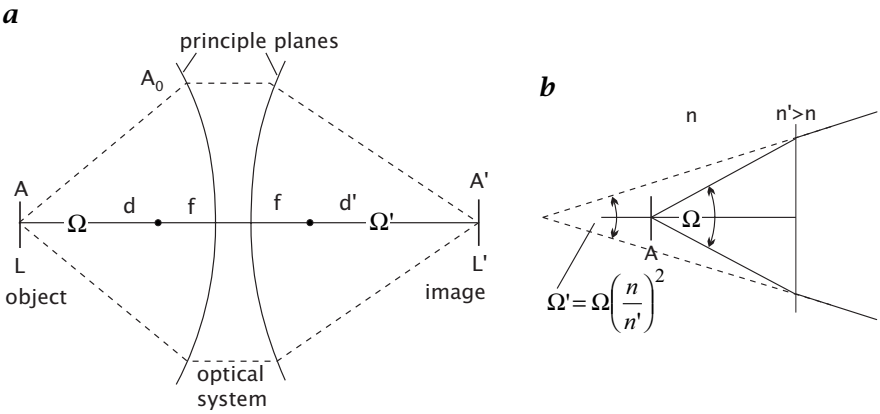

Figure 7.11: *Illustration of radiance invariance:* ***a*** *The product $A\Omega$ is the same in object and image space.* ***b*** *Change of solid angle, when a beam enters an optically denser medium.*

and the following invariance holds:

$$L' = tL \quad \text{for} \quad n' = n. \tag{7.31}$$

The radiance invariance of this form is only valid if the object and image are in media with the same refractive index ($n' = n$). If a beam with radiance L enters a medium with a higher refractive index, the radiance increases as the rays are bent towards the optical axis (Fig. 7.11b). Thus, more generally the ratio of the radiance and the refractive index squared remains invariant:

$$L'/n'^2 = tL/n^2 \tag{7.32}$$

From the radiance invariance, we can immediately infer the irradiance on the image plane to be

$$E' = L'\Omega' = L'\pi \left(\frac{r}{f + d'}\right)^2 = L'\pi \sin^2 \alpha' = tL\pi \sin^2 \alpha'. \tag{7.33}$$

This equation does not consider the fall-off with $\cos^4 \theta$ in Eq. (7.27) because we did not consider oblique principal rays.

Radiance invariance considerably simplifies computation of image irradiance and the propagation of radiation through complex optical systems. Its fundamental importance can be compared to the principles in geometric optics that radiation propagates in such a way that the optical path nd (real path times the index of refraction) takes an extreme value.

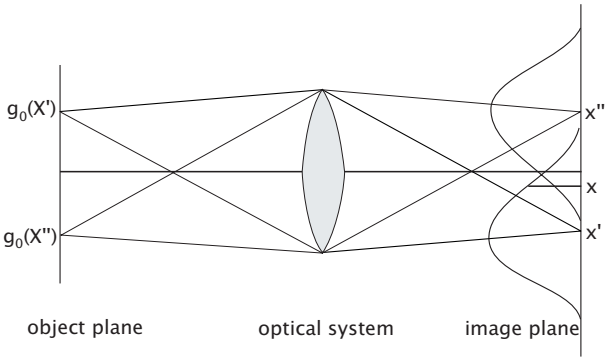

Figure 7.12: *Image formation by convolution with the point spread function $h(x)$. A point at X' in the object plane results in an intensity distribution with a maximum at the corresponding point x' on the image plane. At a point x on the image plane, the contributions from all points x', i.e., $g_i'(x')h(x - x')$, must be integrated.*

7.6 *Linear System Theory of Imaging*

In Section 4.2 we discussed linear shift-invariant filters (convolution operators) as one application of linear system theory. Imaging is another example that can be described with this powerful concept. Here we will discuss optical imaging in terms of the 2-D and 3-D point spread function (Section 7.6.1) and optical transfer function (Section 7.6.2).

7.6.1 *Point Spread Function*

Previously it was seen that a point in the 3-D object space is not imaged onto a point in the image space but onto a more or less extended area with varying intensities. Obviously, the function that describes the imaging of a point is an essential feature of the imaging system and is called the *point spread function*, abbreviated as *PSF*. We assume that the PSF is not dependent on position. Then optical imaging can be treated as a *linear shift-invariant system* (*LSI*) (Section 4.2).

If we know the PSF, we can calculate how any arbitrary 3-D object will be imaged. To perform this operation, we think of the object as decomposed into single points. Figure 7.12 illustrates this process. A point X' at the object plane is projected onto the image plane with an intensity distribution corresponding to the point spread function h. With $g_i'(\boldsymbol{x}')$ we denote the intensity values at the object plane $g_o'(\boldsymbol{X}')$ projected onto the image plane but without any defects through the imaging. Then the intensity of a point $\boldsymbol{x}$ at the image plane is computed by integrating the contributions from the point spread functions which have their maximums at x' (Fig. 7.12):

$$g_i(\boldsymbol{x}) = \int_{-\infty}^{\infty} g_i'(\boldsymbol{x}')h(\boldsymbol{x} - \boldsymbol{x}')\mathrm{d}^2x' = (g_i' * h)(\boldsymbol{x}). \tag{7.34}$$

The operation in Eq. (7.34) is known as a *convolution*. Convolutions play an essential role in image processing. Convolutions are not only involved in image formation but also in many image-processing operations. In case of image formation, a convolution obviously "smears" an image and reduces the resolution.

This effect of convolutions can be most easily demonstrated with image structures that show periodic gray value variations. As long as the repetition length, the *wavelength*, of this structure is larger than the width of the PSF, it will suffer no significant changes. As the wavelength decreases, however, the amplitude of the gray value variations will start to decrease. Fine structures will finally be smeared out to such an extent that they are no longer visible. These considerations emphasize the important role of periodic structures and lead naturally to the introduction of the *Fourier transform* which decomposes an image into the periodic gray value variations it contains (Section 2.3).

Previous considerations showed that formation of a two-dimensional image on the image plane is described entirely by its PSF. In the following we will extend this concept to three dimensions and explicitly calculate the point spread function within the limit of geometric optics, i. e., with a perfect lens system and no diffraction. This approach is motivated by the need to understand three-dimensional imaging, especially in microscopy, i. e., how a point in the 3-D object space is imaged not only onto a 2-D image plane but into a 3-D image space.

First, we consider how a fixed point in the object space is projected into the image space. From Fig. 7.8 we infer that the radius of the unsharpness disk is given by

$$\epsilon_i = \frac{r x_3}{d_i}. \tag{7.35}$$

The index i of ϵ indicates the image space. Then we replace the radius of the aperture r by the maximum angle under which the lens collects light from the point considered and obtain

$$\epsilon_i = \frac{d_o}{d_i} x_3 \tan\alpha. \tag{7.36}$$

This equation gives us the edge of the PSF in the image space. It is a double cone with the x_3 axis in the center. The tips of both the cones meet at the origin. Outside the two cones, the PSF is zero. Inside the cone, we can infer the intensity from the conservation of radiation energy. Since the radius of the cone increases linearly with the distance to the plane of focus, the intensity within the cone decreases quadratically. Thus the PSF $h_i(\boldsymbol{x})$ in the image space is given by

$$\begin{aligned} h_i(\boldsymbol{x}) &= \frac{I_0}{\pi(\frac{d_o}{d_i} x_3 \tan\alpha)^2} \Pi \frac{(x_1^2 + x_2^2)^{1/2}}{2\frac{d_o}{d_i} x_3 \tan\alpha} \\ &= \frac{I_0}{\pi(\frac{d_o}{d_i} z \tan\alpha)^2} \Pi \frac{r}{2\frac{d_o}{d_i} z \tan\alpha}, \end{aligned} \tag{7.37}$$

where I_0 is the light intensity collected by the lens from the point, and Π is the *box function*, which is defined as

$$\Pi(x) = \begin{cases} 1 & |x| \le 1/2 \\ 0 & \text{otherwise} \end{cases}. \tag{7.38}$$

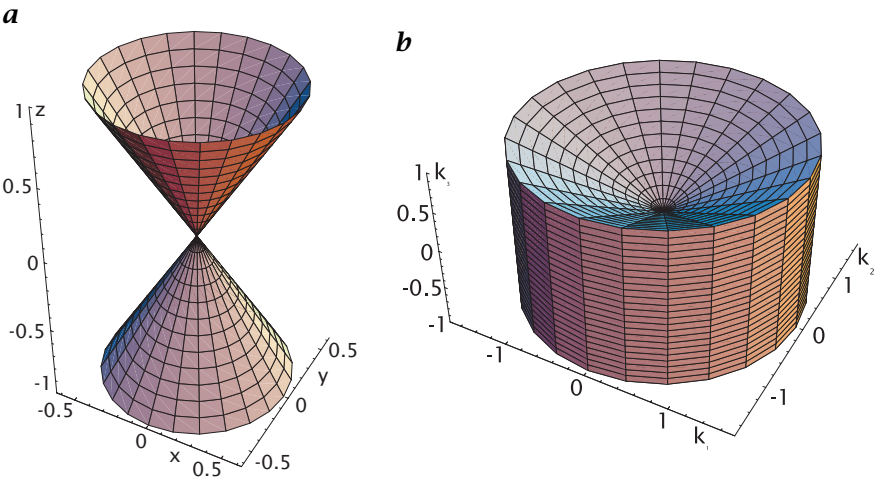

Figure 7.13: ***a*** *3-D PSF and* ***b*** *3-D OTF of optical imaging with a lens, back-projected into the object space. Lens aberrations and diffraction effects are neglected.*

The last expression in Eq. (7.37) is written in cylindrical coordinates (r, ϕ, z) to take into account the circular symmetry of the PSF with respect to the x_3 axis. In a second step, we discuss what the PSF in the image space refers to in the object space, since we are interested in how the effects of the imaging are projected back into the object space. We have to consider both the lateral and axial magnification. First, the image, and thus also ε, are larger than the object by the factor d_i/d_o. Second, we must find the planes in object and image space corresponding to each other. This problem has already been solved in Section 7.4.2. Equation Eq. (7.16) relates the image to the camera coordinates. In effect, the back-projected radius of the unsharpness disk, ϵ_o, is given by

$$\epsilon_o = X_3 \tan\alpha, \tag{7.39}$$

and the PSF, back-projected into the object space, by

$$h_o(\boldsymbol{X}) = \frac{I_0}{\pi(X_3\tan\alpha)^2}\Pi\frac{(X_1^2+X_2^2)^{1/2}}{2X_3\tan\alpha} = \frac{I_0}{\pi(Z\tan\alpha)^2}\Pi\frac{R}{2Z\tan\alpha}. \tag{7.40}$$

The double cone of the PSF, back-projected into the object space, shows the same opening angle as the lens (Fig. 7.13). In essence, $h_0(x)$ in Eq. (7.40) gives the effect of optical imaging disregarding geometric scaling.

7.6.2 *Optical Transfer Function*

Convolution with the PSF in the space domain is a quite complex operation. In Fourier space, however, it is performed as a multiplication of complex numbers. In particular, convolution of the 3-D object $g'_o(\boldsymbol{X})$ with the PSF $h_o(\boldsymbol{X})$ corresponds in Fourier space to a multiplication of the Fourier transformed object $\hat{g}'_o(\boldsymbol{k})$ with the Fourier transformed PSF, the *optical transfer function* or *OTF*

$\hat{h}_o(\boldsymbol{k})$. In this section, we consider the optical transfer function in the object space, i. e., we project the imaged object back into the object space. Then the image formation can be described by:

	Imaged object		Imaging		Object
Space domain	$g_o(\boldsymbol{X})$	=	$h_o(\boldsymbol{X})$	$*$	$g'_o(\boldsymbol{X})$
Fourier domain	$\hat{g}_o(\boldsymbol{k})$	=	$\hat{h}_o(\boldsymbol{k})$	$\cdot$	$\hat{g}'_o(\boldsymbol{k})$.

(7.41)

This correspondence means that we can describe optical imaging with either the point spread function or the optical transfer function. Both descriptions are complete. As with the PSF, the OTF has an illustrative meaning. As the Fourier transform decomposes an object into periodic structures, the OTF tells us how the optical imaging process changes these periodic structures. An OTF of 1 for a particular wavelength means that this periodic structure is not affected at all. If the OTF is 0, it disappears completely. For values between 0 and 1 it is attenuated correspondingly. Since the OTF is generally a complex number, not only the amplitude of a periodic structure can be changed but also its phase.

Direct calculation of the OTF is awkward.

Here several features of the Fourier transform are used, especially its linearity and separability, to decompose the PSF into suitable functions, which can be transformed more easily. Two possibilities are demonstrated. They are also more generally instructive, since they illustrate some important features of the Fourier transform.

The first method for calculating the OTF decomposes the PSF into a bundle of δ lines intersecting at the origin of the coordinate system. They are equally distributed in the cross-section of the double cone. We can think of each δ line as being one light ray. Without further calculations, we know that this decomposition gives the correct quadratic decrease in the PSF, because the same number of δ lines intersect a quadratically increasing area. The Fourier transform of a δ line is a δ plane which is perpendicular to the line ($\succ$ R5). Thus the OTF is composed of a bundle of δ planes. They intersect the k_1k_2 plane at a line through the origin of the k space under an angle of at most α. As the Fourier transform preserves rotational symmetry, the OTF is also circular symmetric with respect to the k_3 axis. The OTF fills the whole Fourier space except for a double cone with an angle of $\pi/2 - \alpha$. In this sector the OTF is zero. The exact values of the OTF in the non-zero part are difficult to obtain with this decomposition method.

We will infer it with another approach, based on the separability of the Fourier transform. We think of the double cone as layers of disks with varying radii which increase with $|x_3|$. In the first step, we perform the Fourier transform only in the x_1x_2 plane. This transformation yields a function with two coordinates in the k space and one in the x space, (k_1, k_2, x_3), respectively (q, φ, z) in cylinder coordinates. Since the PSF Eq. (7.40) depends only on r (rotational symmetry around the z axis), the two-dimensional Fourier transform corresponds to a one-dimensional *Hankel transform* of zero order [13]:

$$\begin{aligned} h(r,z) &= \frac{I_0}{\pi(z\tan\alpha)^2}\Pi\left(\frac{r}{2z\tan\alpha}\right) \\ \check{h}(q,z) &= I_0\frac{\mathrm{J}_1(2\pi zq\tan\alpha)}{\pi zq\tan\alpha}. \end{aligned} \tag{7.42}$$

The Fourier transform of the disk thus results in a function that contains the *Bessel function* J_1 (≻ R5).

As a second step, we perform the missing one-dimensional Fourier transform in the z direction. Equation Eq. (7.42) shows that $\breve{h}(q,z)$ is also a Bessel function in z. This time, however, the Fourier transform is one-dimensional. Thus we obtain not a disk function but a circle function (≻ R5):

$$\frac{\mathrm{J}_1(2\pi x)}{x} \quad \circ\!\!-\!\!\bullet \quad 2\left(1-k^2\right)^{1/2} \Pi\left(\frac{k}{2}\right). \tag{7.43}$$

If we finally apply the Fourier transform *scaling theorem* (≻ R4),

$$\begin{aligned} &\text{if} \quad f(x) \quad \circ\!\!-\!\!\bullet \quad \hat{f}(k), \\ &\text{then} \quad f(ax) \quad \circ\!\!-\!\!\bullet \quad \frac{1}{|a|}\hat{f}\left(\frac{k}{a}\right), \end{aligned} \tag{7.44}$$

we obtain

$$\hat{h}(q,k_3) = \frac{2I_0}{\pi|q\tan\alpha|}\left(1-\frac{k_3^2}{q^2\tan^2\alpha}\right)^{1/2} \Pi\left(\frac{k_3}{2q\tan\alpha}\right). \tag{7.45}$$

A large part of the OTF is zero. This means that spatial structures with the corresponding directions and wavelengths completely disappear. In particular, this is the case for all structures in the z direction, i. e., perpendicular to the image plane. Such structures get completely lost and cannot be reconstructed without additional knowledge.

We can only see 3-D structures if they also contain structures parallel to the image plane. For example, it is possible to resolve points or lines that lie above each other. We can explain this in the x space as well as in the k space. The PSF blurs the points and lines, but they can still be distinguished if they are not too close to each other.

Points or lines are extended objects in Fourier space, i. e., constants or planes. Such extended objects partly coincide with the non-zero parts of the OTF and thus will not vanish entirely. Periodic structures up to an angle of α to the k_1k_2 plane, which just corresponds to the opening angle of the lens, are not eliminated by the OTF. Intuitively, we can say that we are able to recognize all 3-D structures that we can actually look into. All we need is at least one ray that is perpendicular to the wave number of the structure and, thus, run in the direction of constant gray values.

7.6.3 *Diffraction-Limited Optical Systems*

Light is electromagnetic radiation and as such subject to wave-related phenomena. When a parallel bundle of light enters an optical system, it cannot be focused to a point even if all aberrations have been eliminated. Diffraction at the aperture of the optical system blurs the spot at the focus to a size of at least the order of the wavelength of the light. An optical system for which the aberrations have been suppressed to such an extent that it is significantly lower than the effects of diffraction is called *diffraction-limited*.

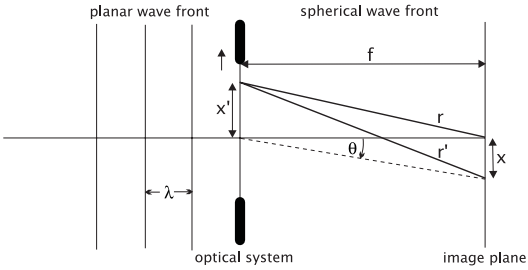

Figure 7.14: *Diffraction of a planar wave front at the aperture stop of an optical system. The optical system converts the incoming planar wave front into spherical wave fronts in all directions converging at the image plane. For further details, see text.*

A rigorous treatment of diffraction according to Maxwell's equations is mathematically quite involved ([12], [39, Chapters 9 and 10], and [85, Chapter 3]). The diffraction of a planar wave at the aperture of lenses, however, can be treated in a simple approximation known as *Fraunhofer diffraction*. It leads to a fundamental relation.

We assume that the aperture of the optical system is pierced by a planar wave front coming from an object at infinity (Fig. 7.14). The effect of a perfect lens is that it bends the planar wave front into a spherical wave front with its origin at the focal point at the optical axis. Diffraction at the finite aperture of the lens causes light also to go in other directions. This effect can be taken into account by applying *Huygens' principle* at the aperture plane. This principle states that each point of the wave front can be taken as the origin of a new in-phase spherical wave. All these waves superimpose at the image plane to form an image of the incoming planar wave. The path lengths from a point $\boldsymbol{x}'$ at the image aperture to the focal point and to a point with an offset $\boldsymbol{x}$ at the image plane (Fig. 7.14) are given by

$$s = \sqrt{x'^2 + y'^2 + f^2} \quad \text{and} \quad s' = \sqrt{(x'-x)^2 + (y'-y)^2 + f^2}, \tag{7.46}$$

respectively. The difference between these two pathes under the condition that $x \ll f$, i. e., neglecting quadratic terms in x and y, yields

$$s' - s \approx -\frac{xx' + yy'}{f}. \tag{7.47}$$

This path difference results in a phase difference of

$$\Delta\varphi = \frac{2\pi(s'-s)}{f\lambda} = -\frac{2\pi(xx'+yy')}{f\lambda} = -\frac{2\pi(\boldsymbol{x}\boldsymbol{x}')}{f\lambda} \tag{7.48}$$

for a wave with the wavelength λ.

Now we assume that $\psi'(x')$ is the amplitude distribution of the wave front at the aperture plane. Note that this is a more general approach than just using a simple box function for an aperture stop. We want to treat the more general case of arbitrarily varying amplitude of the wave front or any type of aperture

functions. If we use a complex-valued $\psi'(x')$, it is also possible to include effects that result in a phase shift in the aperture.

Then the superimposition of all spherical waves $\psi'(x')$ at the image plane with the phase shift given by Eq. (7.48) yields

$$\boxed{\psi(\boldsymbol{x}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi'(\boldsymbol{x}') \exp\left(-2\pi \mathrm{i} \frac{\boldsymbol{x}'\boldsymbol{x}}{f\lambda}\right) \mathrm{d}^2 x'.} \tag{7.49}$$

This equation means that the amplitude and phase distribution $\psi(x)$ at the focal plane is simply the 2-D Fourier transform (see Eq. (2.32)) of the amplitude and phase function $\psi'(x')$ at the aperture plane.

For a *circular aperture*, the amplitude distribution is given by

$$\psi'(\boldsymbol{x}') = \Pi\left(\frac{|\boldsymbol{x}'|}{2r}\right), \tag{7.50}$$

where r is the radius of the aperture. The Fourier transform of Eq. (7.50) is given by the Bessel function of first order (≻ R4):

$$\psi(\boldsymbol{x}) = \psi_0 \frac{I_1(2\pi x r/f\lambda)}{\pi x r/f\lambda}. \tag{7.51}$$

The irradiance E on the image plane is given by the square of the amplitude:

$$E(\boldsymbol{x}) = |\psi(\boldsymbol{x})|^2 = \psi_0^2 \left(\frac{I_1(2\pi x r/f\lambda)}{\pi x r/f\lambda}\right)^2. \tag{7.52}$$

The diffraction pattern has a central spot that contains 83.9 % of the energy and encircling rings with decreasing intensity (Fig. 7.15a). The distance from the center of the disk to the first dark ring is

$$\boxed{\Delta x = 0.61 \cdot \frac{f}{r}\lambda = 1.22\lambda n_f.} \tag{7.53}$$

At this distance, two points can clearly be separated (Fig. 7.15b). This is the *Rayleigh criterion* for resolution of an optical system. The resolution of an optical system can be interpreted in terms of the angular resolution of the incoming planar wave and the spatial resolution at the image plane. Taking the Rayleigh criterion Eq. (7.53), the angular resolution $\Delta\theta_0 = \Delta x/f$ is given as

$$\Delta\theta_0 = 0.61\frac{\lambda}{r}. \tag{7.54}$$

Thus, the angular resolution does not depend at all on the focal length but only the aperture of the optical system in relation to the wavelength of the electromagnetic radiation.

In contrast to the angular resolution, the spatial resolution Δx at the image plane, depends according to Eq. (7.53) only on the relation of the radius of the lens aperture to the distance f of the image of the object from the principal

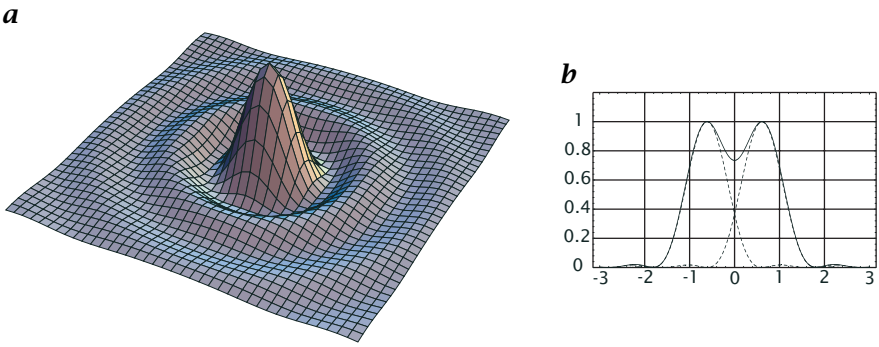

Figure 7.15: ***a*** *Irradiance $E(\boldsymbol{x})$ of the diffraction pattern ("Airy disk") at the focal plane of an optical system with a uniformly illuminated circular aperture according to Eq. (7.52).* ***b*** *Illustration of the resolution of the image of two points at a distance $x/(n_f\lambda) = 1.22$.*

point. Instead of the f-number we can use in Eq. (7.53) the *numerical aperture* which is defined as

$$n_a = n \sin\theta_0 = \frac{2n}{n_f}. \tag{7.55}$$

We assume now that the image-sided index of refraction n may be different from 1. Here θ_0 is the opening angle of the light cone passing from the center of the image plane through the lens aperture. Then

$$\Delta x = 0.61 \frac{\lambda}{n_a'}. \tag{7.56}$$

Therefore, the absolute resolution at the image plane does not at all depend again on the focal length of the system but only the numerical aperture of the image cone.

As the light way can be reversed, the same arguments apply for the object plane. The spatial resolution at the object plane depends only on the numerical aperture of the object cone, i. e., the opening angle of the cone entering the lens aperture:

$$\Delta X = 0.61 \frac{\lambda}{n_a}. \tag{7.57}$$

These simple relations are helpful to evaluate the performance of optical systems. Since the maximum numerical aperture of optical systems is about one, no smaller structures than about half the wavelength can be resolved.

7.7 Homogeneous Coordinates

In computer graphics, the elegant formalism of *homogeneous coordinates* [42, 52, 134] is used to describe all the transformations we have discussed so far, i. e., translation, rotation, and perspective projection, in a unified framework.

This formalism is significant, because the whole image formation process can be expressed by a single 4×4 matrix.

A four-component column vector represents homogeneous coordinates

$$\boldsymbol{X}' = [tX_1', tX_2', tX_3', t]^T, \tag{7.58}$$

from which ordinary three-dimensional coordinates are obtained by dividing the first three components of the homogeneous coordinates by the fourth. Any arbitrary transformation can be obtained by premultiplying the homogeneous coordinates with a 4×4 matrix $\boldsymbol{M}$. In particular, we can obtain the image coordinates

$$\boldsymbol{x} = [sx_1, sx_2, sx_3, s]^T \tag{7.59}$$

by

$$\boldsymbol{x} = \boldsymbol{M}\boldsymbol{X}. \tag{7.60}$$

As matrix multiplication is associative, we can view the matrix $\boldsymbol{M}$ as composed of many transformation matrices, performing such elementary transformations as *translation*, *rotation* around a coordinate axis, *perspective projection*, and *scaling*. The transformation matrices for the elementary transforms are readily derived:

$$\begin{aligned}
\boldsymbol{T} &= \begin{bmatrix} 1 & 0 & 0 & T_1 \\ 0 & 1 & 0 & T_2 \\ 0 & 0 & 1 & T_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} && \text{Translation by } [T_1, T_2, T_3]^T \\
\boldsymbol{R}_{x_1} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} && \text{Rotation about } X_1 \text{ axis by } \theta \\
\boldsymbol{R}_{x_2} &= \begin{bmatrix} \cos\phi & 0 & \sin\phi & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\phi & 0 & \cos\phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} && \text{Rotation about } X_2 \text{ axis by } \phi \\
\boldsymbol{R}_{x_3} &= \begin{bmatrix} \cos\psi & -\sin\psi & 0 & 0 \\ \sin\psi & \cos\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} && \text{Rotation about } X_3 \text{ axis by } \psi \\
\boldsymbol{S} &= \begin{bmatrix} s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \\ 0 & 0 & s_3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} && \text{Scaling} \\
\boldsymbol{P} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/d' & 1 \end{bmatrix} && \text{Perspective projection.}
\end{aligned} \tag{7.61}$$

Perspective projection is formulated slightly differently from the definition in Eq. (7.11). Premultiplication of the homogeneous vector

$$\boldsymbol{X} = [tX_1, tX_2, tX_3, t]^T$$

with $\boldsymbol{P}$ yields

$$\left[tX_1, tX_2, tX_3, t\frac{d' - X_3}{d'}\right]^T, \tag{7.62}$$

from which we obtain the image coordinates by division through the fourth coordinate

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} X_1\dfrac{d'}{d' - X_3} \\ X_2\dfrac{d'}{d' - X_3} \end{bmatrix}. \tag{7.63}$$

From this equation we can see that the image plane is positioned at the origin, since if $X_3 = 0$, both image and world coordinates are identical. The center of projection has been shifted to $[0, 0, -d']^T$.

Complete transformations from world coordinates to image coordinates can be composed of these elementary matrices. Strat [196], for example, proposed the following decomposition:

$$\boldsymbol{M} = \boldsymbol{CSPR}_z\boldsymbol{R}_y\boldsymbol{R}_x\boldsymbol{T}. \tag{7.64}$$

The scaling $\boldsymbol{S}$ and cropping (translation) $\boldsymbol{C}$ are transformations taking place in the two-dimensional image plane. Strat [196] shows how the complete transformation parameters from camera to world coordinates can be determined in a noniterative way from a set of calibration points whose positions in the space are exactly known. In this way, an absolute calibration of the outer camera parameters position and orientation and the inner parameters piercing point of the optical axis, focal length, and pixel size can be obtained.

7.8 Exercises

7.1: ****Imaging with a pinhole camera**

1. What is the relation between object and image coordinates for a pinhole camera?
2. What geometric object is the image of a straight line with the points $\boldsymbol{A}$ and $\boldsymbol{B}$, a triangle with the points $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$, and of a planar and nonplanar quadrangle?
3. Assume that you know the length of the straight line and the position $\boldsymbol{A}$ of one of the end points in world coordinates. Is it then possible to determine the second end point $\boldsymbol{B}$ from the image coordinates $\boldsymbol{a}$ and $\boldsymbol{b}$?

7.2: ***Geometry of imaging with x-rays**

Can the imaging with penetrating x-rays that emerge from a single point and are measured at a projection screen also be described by projective imaging? The object is now located *between* the x-ray source and the projection screen. How is the relation between image and world coordinates in this case? Prepare a sketch of the geometry.

7.3: ***Depth of field with x-ray imaging**

Is it possible to limit the depth of field with x-ray imaging? Hint: You cannot use any lens with x-rays. The depth of field is related to the fact that the lens collects rays from a point of the object that are going into a range of directions. How can this principle be used with a non-imaging system? The object to be inspected does not move.

7.4: *__High depth of field__

You are facing the following problem. An object should be measured with the maximum possible depth of field. The illumination conditions, which you cannot change, limit the aperture n_f to a maximum value of 4. The object has an extension of $320 \times 240\,\mathrm{mm}^2$ and must fill the whole image size when imaged from a distance of $2.0 \pm 0.5\,\mathrm{m}$. Two cameras with a resolution of 640×480 pixels are at your disposal. The pixel size of one camera is $9.9 \times 9.9\,\mu\mathrm{m}^2$, that of the other camera $5.6 \times 5.6\,\mu\mathrm{m}^2$ (≻ R2). You can use any focal length f of the lens. Questions:

1. Which focal length do you select?
2. Which of the two cameras delivers the larger depth of field?

7.5: *__Diffraction-limited resolution__

At which aperture n_f is the diffraction-limited resolution equal to the size of the sensor element? Use $4.4 \times 4.4\,\mu\mathrm{m}^2$ and $6.7 \times 6.7\,\mu\mathrm{m}^2$ large sensor elements. What happens at larger n_f values?

7.9 *Further Readings*

In this chapter, only the basic principles of imaging techniques are discussed. A more detailed discussion can be found in Jähne [89] or Richards [165]. The geometrical aspects of imaging are also of importance for computer graphics and are therefore treated in detail in standard textbooks on computer graphics, e. g. Watt [211] or Foley et al. [52]. More details about optical engineering can be found in the following textbooks: Iizuka [85] (especially about Fourier optics) and Smith [191]. Riedl [168] focuses on the design of infrared optics. In this chapter, the importance of linear system theory has been stressed for the description of an optical system. Linear system theory has widespread applications throughout science and engineering, see, e. g., Close and Frederick [24] or Dorf and Bishop [36].