AMERICAN JOURNAL of PHYSICS

Pinhole Imagery

M. Young

Citation: American Journal of Physics **40**, 715 (1972); doi: 10.1119/1.1986624
View online: http://dx.doi.org/10.1119/1.1986624
View Table of Contents: http://scitation.aip.org/content/aapt/journal/ajp/40/5?ver=pdfcov
Published by the American Association of Physics Teachers

Pinhole Imagery

M. YOUNG
Electrophysics Division
Rensselaer Polytechnic Institute
Troy, New York 12181
(Received 6 December 1971; revised 27 January 1972)

Despite its antiquity and apparent simplicity, the pinhole camera can be an interesting and useful device, boasting freedom from linear distortion, infinite depth of field, and wide angular field. The pinhole camera may be optimized for best resolution, and even its low aperture is largely offset by the availability of modern light sources and fast detectors. Clear understanding of this simple optical system yields a good deal of insight into physical optics and optical instruments.

INTRODUCTION

One of the very earliest optical instruments was no doubt the camera obscura or pinhole camera. The invention of the pinhole camera is generally credited to della Porta in approximately 1600. Mach believes the instrument to have been used earlier by Leonardo da Vinci during his studies on perspective,[1] and the pinhole camera was likely known to the ancient Greeks as well. Despite this antiquity, the pinhole camera still offers several advantages over more sophisticated optics, mainly in applications where resolution is not a major factor.[2] Pinhole optics boasts complete freedom from linear distortion, virtually infinite depth of field, and a very wide angular field. The pinhole's light gathering ability is largely offset by the availability of modern light sources and high-speed films and video detectors. In addition, pinhole optics are readily available at wavelengths where other optics are not.

Finally, we can learn a great deal of practical physics from a study of this simple device. The pinhole camera is a useful tool for illustrating such concepts as limit of resolution or the relationship between ray and wave approximations.

Suitable pinholes are very easy to make. For example, I once had occasion to make an array of 25 μm pinholes (for another purpose). I punched the pinholes into 50 μm brass shim stock by placing a hand sewing needle in a milling machine and using the vertical feed on the machine to force the needle through the brass and into a freshly faced-off block of lead. Punched carefully, "deburred," reamed with a needle point, and cleaned, the holes are repeatably the required size. Larger pinholes can be punched by hand.

Selwyn's paper on the pinhole camera is an excellent work in which the author shows (theoretically, but with virtually no mathematics) that it is in a sense necessary to focus a pinhole camera, and that the camera suffers from certain aberrations.[3] Selwyn derives an "optimum" size for the pinhole by finding the relative irradiance at the center of the image of a point. Martin assumes the "best" pinhole to include a single Fresnel zone,[4] while others[5,6] take an approach similar to the one I shall present below. These different treatments yield substantially the same result.

In what follows, I shall outline a simple theory for the pinhole camera, study the pinhole as an optical device, and describe some experimental results. Close examination of this simple instrument yields a good deal of insight not only into optics, but into the basic physics as well.

THEORY

Point Images

An image of a distant point, such as a star, is never itself a point, but rather a small spot. We can think of an image of an extended object as a collection of such spots or point images. Each point image represents approximately the finest detail that can be discerned in the extended image.

A pinhole camera consists of a small hole in an opaque screen. Each object point casts a shadow of this screen onto a viewing screen. The shadow

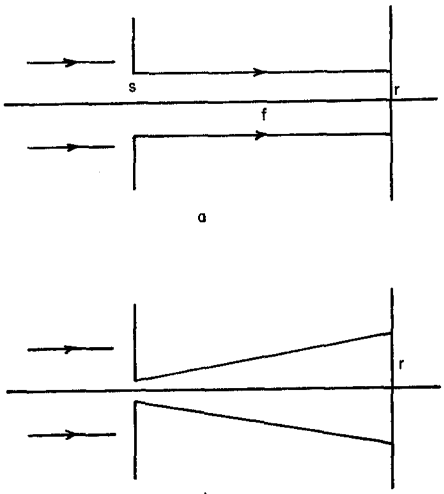

FIG. 1. Pinhole camera "focused" at infinity: (a) large pinhole, ray approximation; (b) small pinhole, Fraunhofer approximation.

of a single bright point is a bright spot, and we call this spot the image of that point. The smaller the point image, the finer the detail that can be observed in an extended object. Thus in many ways the best pinhole is the one which produces the smallest point image.

When the pinhole radius is "large" the image of a distant point is found by applying geometrical or ray optics. It is a uniform disk, which is no more than the geometrical shadow of the screen containing the pinhole. In the notation of Fig. 1, the pinhole radius is s and the image radius r. In ray approximation clearly $r=s$ for a very distant object point.

If we reduce the pinhole size in an effort to make the point image smaller, we will be successful only to a limited extent. When we decrease s beyond a certain value (depending on the wavelength λ and the "focal length" f) we will find that the image is no longer a uniform disk. Rather, it shows a considerable ringlike structure indicating that geometrical optics no longer provides an adequate description. It is thus necessary to take the wave nature of light into account.

If we reduce the pinhole size still further, we find that the image radius no longer decreases but begins to increase. At some point, Fraunhofer or far-field diffraction becomes applicable, and r is inversely proportional to s. The image of a distant point is just the diffraction pattern of the pinhole. For a circular pinhole, its radius is given by $r=0.61\lambda f/s$.[7]

Thus, the two limiting cases (of a large and small pinhole) give rise to a large point image. We can estimate the "optimum" pinhole size by plotting r vs s as in Fig. 2. For large s, the curve is a straight line with unit slope. For small s, it is a hyperbola. Where the curves cross, neither approximation is valid, but it is nevertheless clear that the smallest image occurs somewhere in this region. The curves cross where geometrical and Fraunhofer approximations give the same result, $s^2 \sim \lambda f$, or, for given s, $f \sim s^2/\lambda$.

This value of s is approximately the optimum pinhole radius, and there has been much theoretical work devoted to refining this estimate. Such refinement is essentially impossible because neither approximation is valid in the region of greatest interest, namely, pinholes of approximately the optimum size. There is also conceptual difficulty comparing uniform circular disks with Fraunhofer diffraction patterns. The precise behavior of the pinhole camera near optimum resolution should therefore be determined experimentally.

Two-Point Resolution

We are generally more interested in resolution between points or lines than in isolated point images. When this is so, it is more useful to speak of resolution limit than of image radius r.

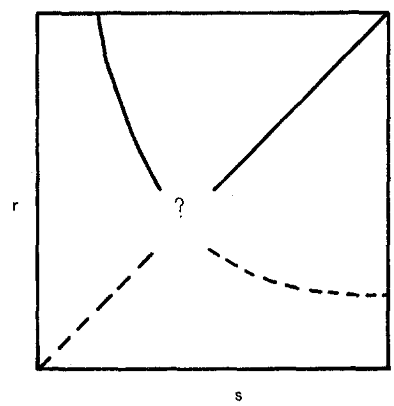

FIG. 2. Image radius r vs pinhole radius s for given focal length f. Solid curves represent limiting cases of high and low aperture.

When the pinhole is large compared with $(\lambda f)^{1/2}$, the image of a single point is a uniformly illuminated disk with radius $r=s$. When two disks are too close together, it is not possible to determine that the object consists of two discrete points. We would probably agree that the points in Fig. 3(a) are not resolvable; here the images are separated by a distance equal to their radius. In Fig. 3(b) the points are separated by twice their radius and seem to be well resolved.

The limit of resolution refers to the separation at which the points are "just" resolved. It is likely somewhere between s and $2s$, and looking ahead to the experimental results, we adopt the criterion that the two point resolution limit is $1.5s$ in geometrical optics approximation.

When the pinhole is small compared to the optimum, Fraunhofer diffraction applies. The image of a point is not a uniform disk, but consists of a strong central maximum surrounded by weak rings. Even within the central maximum, the light is not uniformly distributed, but falls off rapidly away from the center. The preceding discussion is not relevant, and we must use the familiar Rayleigh criterion.[7] The two point resolution limit is thus $0.61\lambda f/s$ in Fraunhofer approximation.

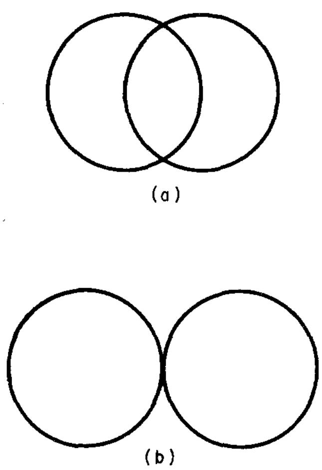

FIG. 3. Limit of resolution in ray approximation. Images are uniform circular disks: (a) unresolved; (b) well resolved.

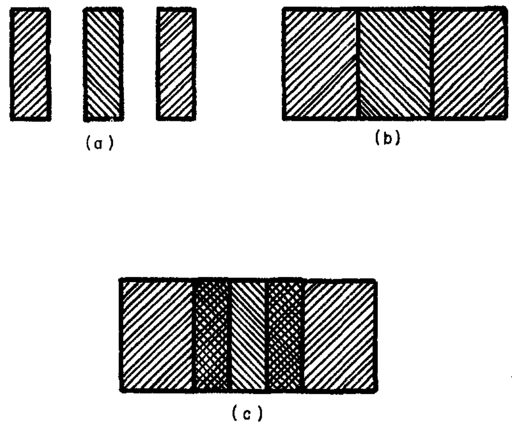

FIG. 4. Images of bar object: (a) good resolution; (b) unresolved; (c) spurious resolution caused by overlapping of adjacent bars.

Spurious Resolution

This is a curious phenomenon, in that a periodic object which is below the limit of resolution can nevertheless exhibit a spurious periodic structure that is easily mistaken for detail. Spurious resolution can be explained by geometrical optics. In the pinhole camera, it occurs only when the pinhole is large enough that ray optics is a good approximation. This suggests that spurious resolution may be a purely geometrical effect.

Figure 4 shows the origin of spurious resolution when the object is a three-bar target. In Fig. 4(a) the resolution limit is small and the bars are well resolved. As we increase the resolution limit, for example by widening the pinhole, we come to the point where the bars are unresolved as in Fig. 4(b). Still widening the pinhole, we find the blurred unresolved bar images beginning to overlap, as in Fig. 4(c).

When the resolution limit is large, there is considerable overlap, and this gives rise to the spurious two-bar structure indicated by the crosshatched region.

For a simple three-bar object, we can easily distinguish resolution from spurious resolution—with spurious resolution three bars appear as two. In a more complex object, spurious resolution can be mistaken for true structure, and it is evident that spurious resolution will often have to be avoided. (Incidentally, spurious resolution is

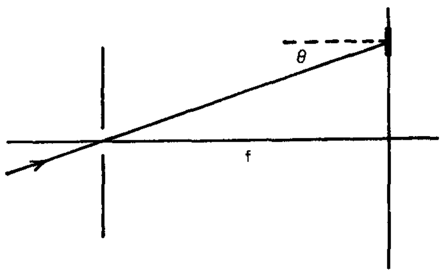

FIG. 5. Object off axis at angle θ. Obliquity gives rise to reduction of exposure by $\cos^4\theta$.

easy to demonstrate by putting a bar object in a slide projector and defocusing the image on the screen. A focusing error so great that the bars cannot be resolved signals the onset of spurious resolution. Defocus a bit more!)

Defects in the Image

Most lens systems suffer from linear distortion. A square is imaged not as a square, but, for example, as a barrel with slightly bulging sides. Simple geometrical considerations show that the pinhole camera presents distortion-free images. Selwyn uses a simple physical argument to show that the wave theory likewise predicts imaging without distortion. The pinhole is the only image-forming device that shows absolutely no distortion over a very wide field.

The pinhole does suffer from a number of aberrations. The optimum focal length for a given aperture is of the order of s^2/λ. In the visible region of the spectrum, wavelength varies from 400–700 nm, or $\pm 20\%$ from the mean wavelength of 550 nm. We thus expect the pinhole camera to exhibit chromatic aberration (unless the pinhole is very large). Fortunately, the effect of this aberration on resolution is found from data presented later to be small.

An important off-axis aberration is astigmatism. When the pinhole is viewed from a point well off the optical axis, it appears elliptical, rather than circular. This can give rise to a serious degradation of the image, although I have shown in Ref. 2 how it is possible to "play" wave optics against ray optics and partially correct this astigmatism.

The light-gathering power of the pinhole camera is low. If the focal length of the camera is made equal to s^2/λ, then the focal ratio or F number for visible light is around 200 for a focal length of a few centimeters. Photographers will regard this as an extremely slow system! Because of its low aperture, however, the pinhole camera offers nearly infinite depth of field.

Our ability to expose very wide angle photographs depends on our ability to accept (or correct for) some loss of exposure in the corners of the image plane. This problem is not unique to the pinhole camera, but affects nearly all imaging systems equally. Suppose one point is imaged on axis and one off axis at angle θ (Fig. 5). From the image plane, the pinhole appears as a bright spot of light. The off-axis image is further from the pinhole by $1/\cos\theta$, and the inverse-square law implies that the exposure will be less there by $1/\cos^2\theta$. In addition, the pinhole appears smaller by $\cos\theta$. The light also falls obliquely off axis and hence illuminates an area larger by $1/\cos\theta$, reducing the exposure by another factor of $\cos\theta$. Combining these three effects we find that the exposure is reduced off axis by a factor of $\cos^4\theta$. If we wish to cover a 90° field, $\theta = 45°$, $\cos^4\theta = 1/4$, and we have to accept two f-stops loss of exposure far off axis. Use of a curved film plane centered on the pinhole will reduce the $\cos^4$ law to a simple cos law and allow wide angle photography with relatively little loss of exposure.

EXPERIMENT

Figure 6 depicts the experimental set-up used in Ref. 2 to study the pinhole camera. The light source is a 650-W home movie lamp that can be purchased in any camera store. To reduce stray light, it had to be enclosed in a metal box and

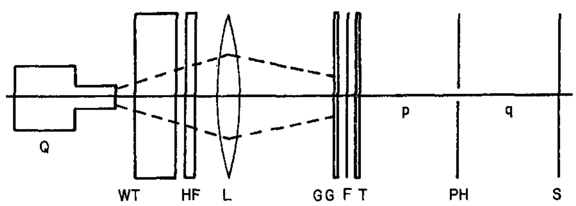

FIG. 6. Experimental pinhole camera. Quartz iodine lamp Q, with heat absorbers WT and HF, illuminates target T through ground glass GG and colored filter F. Pinhole PH images target T onto screen S.

cooled with forced air. The lamp is very hot and will set cardboard afire at a distance of 2 ft: hence the heat absorbing distilled water tank and filter. The lamp illuminated the target through a ground glass and a gelatin filter that provided more or less monochromatic light at 500 nm.

Most of the data were taken with the camera working at unit magnification. The results are presented in Fig. 7 as if the object had been at infinity, and the means for making this transformation are outlined in Ref. 2. Data are plotted in a different form from that of Fig. 2. The resolution limit is expressed in units of pinhole radius s and the focal length in units of s^2/λ. This normalization makes the results pertain to all possible combinations of f, λ, and s.

If we call normalized resolution limit R and normalized focal length ϕ $(=f\lambda/s^2)$, then the solid curves of Fig. 6 represent the limiting values

$$R=1.5$$

and

$$R=0.61\phi$$

for the high and low aperture cases, respectively.

The solid circles show the experimental data, which were actually taken with three different pinholes (160, 420, and 760 μm in diam) under a variety of conditions.[2] Agreement with the simple

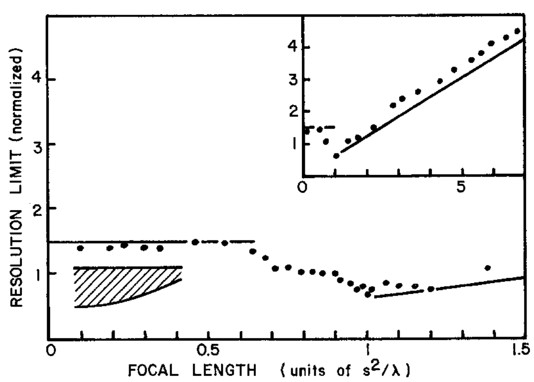

FIG. 7. Resolution limit in units of s vs focal length in units of s^2/λ. Resolution limit is least when $f=s^2/\lambda$ and pinhole occupies single Fresnel zone. Spurious resolution is evident when normalized focal length is less than 0.4 (cross-hatched region).

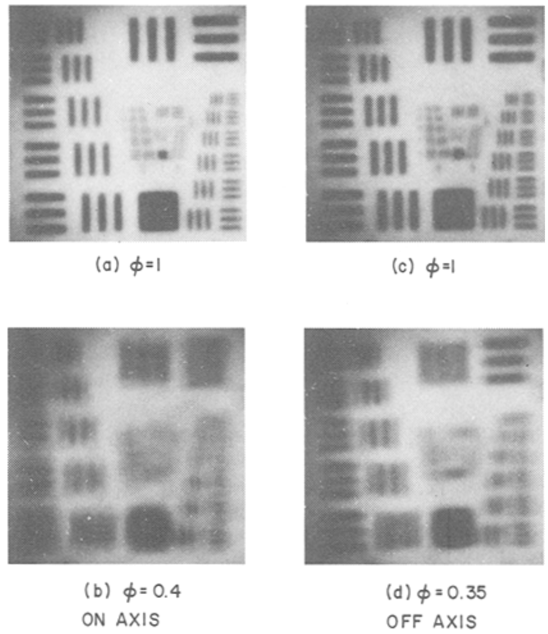

FIG. 8. Resolution target photographed with pinhole camera. Largest bars have frequency of 1.0 l/mm. Upper photographs were exposed with optimum pinhole size, lower, several times larger. Note spurious resolution and astigmatism.

theory is quite good over most of the range. The resolution limit is actually somewhat smaller than the pinhole when f is about equal to s^2/λ. Excursions of 10% or 20% from this value are not excessive, and this explains why chromatic aberration is not as important a factor as we might have expected.

An apparent focusing effect as f approaches s^2/λ also shows up in the data. This effect is not surprising if we think of the pinhole as a Fresnel zone plate[4,7] with a single zone. The focal length of a zone plate is just equal to s^2/λ, and we conclude that the pinhole camera works best when the "focal length" of the camera is just equal to the focal length of the pinhole itself.

Figure 8 shows some photographs of a resolution target exposed with a pinhole camera. Figures 8(a) and 8(b) were exposed on axis, and 8(c) and 8(d) off axis to show the effect of astigmatism. We can easily see spurious resolution in Fig. 8(b), where the pinhole is large, while in Fig. 8(d) astigmatism causes both resolution and spurious resolution to occur for identical sets of bars!

FIG. 9. Snapshot taken with pinhole camera to show characteristic soft focus and great depth of field.

Finally, I used the 420 μm pinhole to "make" a pinhole camera out of a 35 mm camera body and a set of extension tubes. The camera focal length is 90 mm, very nearly the optimum for visible light. Although the camera works at $f/200$, it is quite easy to hand hold the camera out of doors (on a cloudy day!) with a moderately fast film such as Kodak Tri-X. Figure 9 is a snapshot taken with this camera and shows the soft focus and great depth of field of which the pinhole camera is capable.

CONCLUSIONS

The pinhole camera can be a useful instrument for certain specialized applications. Modern light sources and fast detectors make it a practical tool where depth of field, wide angle and freedom from distortion are important considerations, but where resolution is not. The camera gives best performance when its focal length is about equal to s^2/λ. If the pinhole is too large, spurious resolution will be a factor.

Finally, the pinhole camera illustrates very clearly the relation between ray and wave optics. No matter what the size of the pinhole, the light striking it first propagates through the pinhole as if geometrical theory were applicable. This is made clear by the data in Fig. 7. Only when the light has propagated a distance that is significant compared with s^2/λ does the wave theory become important. At about s^2/λ, the light is actually focused somewhat by diffraction. By the time the light has propagated a distance greater than s^2/λ, it has acquired a beam divergence of $0.61\lambda f/s$, and only then is the approximation of Fraunhofer diffraction legitimate.

[1] E. Mach, *The Principles of Physical Optics* (Dover, New York, 1926).

[2] M. Young, Appl. Opt. **10,** 2763 (1971); and references therein, especially Refs. 2–6.

[3] E. W. H. Selwyn, Phot. J. **90B,** 47 (1950).

[4] L. C. Martin, *Technical Optics* (Sir Isaac Pitman and Sons, London, 1959), Vol. 1, pp. 90–91.

[5] A. C. Hardy and F. Perrin, *The Principles of Optics* (McGraw-Hill, New York, 1932), pp. 124–126.

[6] R. Kingslake, *Lenses in Photography* (A. S. Barnes, New York, 1963), revised ed., pp. 60–62.

[7] See, e.g., R. S. Longhurst, *Geometrical and Physical Optics* (Wiley, New York, 1967), 2nd ed.