Pinhole Optics

M. Young

The pinhole camera may often be overlooked because of its apparent simplicity. However, it is a useful and practical device which offers freedom from distortion and virtually infinite depth of field. Its astigmatism can be corrected by proper choice of aperture, and its angular field can be made to exceed 90°. With modern light sources, films, and detectors, even the pinhole's low aperture need not be an important limitation.

Introduction

The first optical instrument, other than the window and plane mirror, was no doubt the camera obscura, or pinhole camera, which was evidently known before the publication of della Porta *ca.* 1600.[1] The pinhole camera still has a number of advantages over more sophisticated optics in applications where resolution is not a major factor. It may often be overlooked because of its apparent simplicity. Nevertheless, the pinhole camera offers complete freedom from linear distortion, virtually infinite depth of field, and an angular field that can be made to exceed 90°. With the high sensitivity of modern film and detectors, even the pinhole's low aperture (say, $f/200$) need not be an important limitation.

Recently, the pinhole camera has seen a number of applications. Newman and Rible reported a pinhole array camera for use in manufacturing integrated circuits.[2] Their 8×8 array of pinholes was designed to eliminate the step-and-repeat camera and to produce a 64-image mask with a single exposure. They recorded lines as fine as 12.5 μm. Gallas and co-workers used a pinhole television camera for a flight simulator which required both large depth of field and large angular field.[3,4] To attain clarity from 7.6 cm to 7.6 m, they chose pinhole optics over a more conventional automatic focusing system. Bernstein and Hai developed an x-ray pinhole camera for use where other optics was unavailable.[5] They recorded time-resolved photographs of the x-ray emission from a plasma device. Waddell has cited additional applications of pinhole optics, including a high speed camera built around a rotating drum perforated with pinholes, a photographically coded flight recorder, and a ballistic attitude recorder for use in rocketry.[6] Finally, I have found the pinhole camera to be a useful teaching tool in explaining and illustrating concepts such as limit of resolution, as well as the relationship between ray approximation and physical optics.

Pinholes suitable for such optical applications are very easily etched[2] or made with a punch[4] or hand sewing needle. For another purpose, I once made a precision array of 25-μm pinholes in thin brass shim stock by putting a needle in a milling machine and using the feeds on the machine to position the brass and to punch the holes against a freshly faced-off block of lead. When the holes had been deburred, carefully reamed with a needle point, and cleaned, they turned out to be accurately the required size.

An over-all look at the pinhole camera is found in an excellent paper by Selwyn,[7] who shows (with virtually no mathematics) that light passing through the pinhole should in a sense be focused, but that the pinhole camera suffers from certain aberrations. He derives the optimum size for the pinhole by finding the relative irradiance at the center of the image of a point. Martin assumes that the best pinhole includes a single Fresnel zone and arrives at substantially the same result as Selwyn.[8] Others[9,10] consider the limiting cases of large and small pinholes, using ray optics and Fraunhofer diffraction, respectively, and in this way estimate the optimum pinhole diameter.

The pinhole camera has also been used to illustrate new theoretical techniques.[11] In addition, Swing and Rooney have derived a general transfer function for the pinhole camera operating on axis and presented results to substantiate their work.[12]

The purpose of this paper is to extend the earlier work to include a wide range of focal lengths and off-axis operation, as well as to review some of the important results of others.

The author is with the Electrophysics Division, Rensselaer Polytechnic Institute, Troy, New York 12181.

Received 18 June 1971.

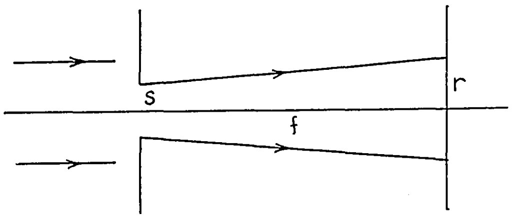


Fig. 1. Pinhole camera focused on a distant point: f is the focal length, s the pinhole radius, and r the radius of the image.

Theory

Point Images

In this paper, we apply a simple theory based on examination of the two limiting cases of a *large* and a *small* pinhole. The image of a point is taken to be a spot, whose radius we call r. The image of an extended object is a collection of point images, each of which represents approximately the limit of resolution. A standard approach to pinhole imagery is to consider a point object at infinity, as shown in Fig. 1. A hole is cut into an opaque screen and a shadow cast onto a viewing screen a distance f away. The bright spot in the center of the shadow is the image of the point, and f may be called the *focal length* of the pinhole camera.

In some ways the best pinhole is the one that produces the smallest point image. When the pinhole radius s is *large*, the image is a uniform disk which is just the geometrical shadow of the pinhole. Thus, the radius r of the image is equal to the pinhole radius: $r = s$. On the other hand, when the pinhole is *small*, the image shows a significant ringlike structure, indicating that geometrical optics does not pertain. The image, which must be described by physical optics, is either a Fresnel or a Fraunhofer diffraction pattern of the pinhole. When Fraunhofer diffraction pertains, r is inversely proportional to s. For a circular hole, $r = 0.61\lambda f/s$. The optimum pinhole size will be a compromise between the large spot produced by a large hole and that produced by Fraunhofer diffraction from a small hole. The smallest image thus occurs roughly where the geometrical optics and Fraunhofer diffraction approximations give the same result, $s^2 \sim 0.61\lambda f$.

As Ref. 3 shows, there has been much theoretical work devoted to finding a more precise value for the optimum pinhole size. The arguments used are basically similar to that in the preceding paragraph. Because they relate uniform geometrical shadows with Fraunhofer diffraction patterns, these arguments are only approximate in nature. The precise pinhole size that yields best resolution should therefore be determined experimentally.

Finite Conjugates

In work with a pinhole camera for certain applications we may well be interested in its performance when the object point is at a finite distance p as in Fig. 2. In this case, we call the image distance q and *define* the operating focal length of the camera by $1/f = (1/p) + (1/q)$. It is further convenient to define the radius g of the geometrical shadow and to measure focal lengths in terms of the far-field distance $f' = s^2/\lambda$. (Note the distinction between the operating focal length f of the camera and s^2/λ, which could be called the focal length of the pinhole.) Applying the same reasoning as before, we find that when the hole is small ($p,q > s^2/\lambda$), the image size r can be expressed in normalized fashion as $\rho = r/g = 0.61\phi$, where ρ is the image radius in units of g and $\phi = f/f'$ is the focal length in units of s^2/λ. Similarly, in geometrical optics approximation ($p,q \ll s^2/\lambda$), $r = g$, or $\rho = 1$, independent of ϕ. If we were to plot ρ against ϕ in both approximations, we would find the transition between the geometrical and physical optics approximations to occur near $f = s^2/\lambda$, as we might have guessed, even though the quantity defined as f is not obviously the focal length of the camera. In view of this definition of f, however, we can apply all the standard formulas of lens optics (magnification, effective f/number, etc.) to pinhole optics.

A more rigorous analysis of the pinhole camera can can be made using the Fresnel integrals,[13] which describe the behavior for all f, but in particular for $f \sim s^2/\lambda$. Although I shall not demonstrate this, it is not difficult to use Fresnel's integrals to show that the distribution of light in the image is a function only of ϕ, as suggested directly above. Thus, by expressing image size and focal length in normalized variables, we are able to describe all possibilities with relatively few observations.

When resolution between point or line objects is important, the resolution limit is generally taken as the radius r of the point image. For the case $\phi > 1$ this is the familiar Rayleigh criterion. When $\phi \ll 1$ (geometrical optics), the light is distributed uniformly throughout the point image, and this criterion does not necessarily apply. Looking ahead to the experiment, we shall take the resolution limit to be $1.5g$ (not g) in geometrical approximation.

Aberrations

One of the advantages of pinhole optics over lens optics is freedom from linear distortion. On the basis

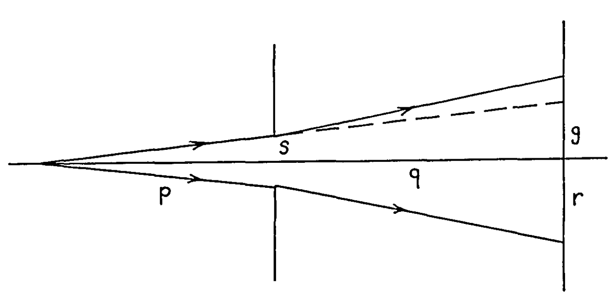


Fig. 2. Pinhole camera focused on a nearby point: p and q are object and image distances, s the pinhole radius, and g the diameter of the geometrical shadow.

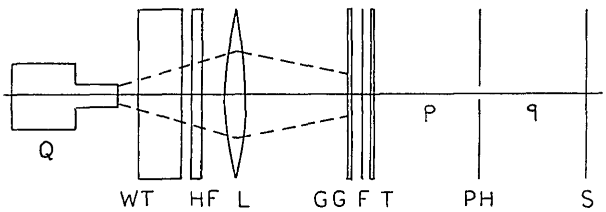


Fig. 3. Experimental pinhole camera. Quartz–iodine lamp Q is filtered by water tank WT and heat-absorbing filter HF. Beam is condensed by lens L onto ground glass GG, which illuminates target T through filter F. Pinhole PH casts image onto screen S. Conjugates are p and q.

of geometrical or ray optics, the pinhole is obviously free from distortion, and Selwyn has used a simple argument based on Fermat's principle to show that the wave theory also predicts distortion-free imagery. On the other hand, the pinhole does suffer from chromatic aberration. To see this, we have only to note that the optimum focal length for a given aperture is on the order of s^2/λ, and that wavelength varies (in the visible) by $\pm 20\%$. It is a simple matter to estimate the amount of transverse chromatic aberration from data presented later, and the effect is small.

Another aberration is a form of astigmatism arising in the image of an off-axis object point. This astigmatism comes about because the pinhole appears elliptical when viewed far enough off axis. Since the optimum focal length in one plane then differs from that in the perpendicular plane, we can also think of this effect as a kind of field curvature.

Choice of pinhole size can, however, reduce aberration. Imagine a point object to be moved in a horizontal plane until it is off axis by an angle a. Then the pinhole appears to have horizontal dimension $s \cos a$, while its vertical dimension remains s. Where geometrical approximation is valid, resolution along a horizontal line is improved, while it is worsened where Fraunhofer diffraction is appropriate. Thus, for a given angle and focal length there ought to be a pinhole size for which there is no astigmatism. As we shall see, this pinhole size satisfies the relation $(s \cos a)^2/\lambda < f < s^2/\lambda$. Careful choice of pinhole size can produce an image with little astigmatism over a very large angular field.

Photometry

The irradiance (or illuminance) in the image plane of the pinhole camera will depend on the f/number $F = f/2s$, just as in an ordinary camera. [If the pinhole is operating with high magnification $m = q/p$, then F is replaced by the usual effective f/number, $F(1 + m)$.] Further, it is easy to see that the familiar $\cos^4$ law applies to the pinhole for the same reasons as to the lens. For an extended object the irradiance at angle a off axis is less by a factor of $\cos^4 a$ than the irradiance on axis. The angular field that we can use may depend on our ability to tolerate or to correct a large variation in exposure between the center and the edge of the image plane. For example, if we accept a difference of two f/stops, $\cos^4 a = \frac{1}{4}$, and the total allowable field is $2a = 90°$. A curved film plane reduces the $\cos^4$ law to a simple cos law and may allow fields wider than 90°.

Experiment

Figure 3 shows an arrangement used to study pinhole optics. The light source Q is a 650-W quartz–iodine lamp of the sort used for amateur movies. The lamp was enclosed in a metal box and cooled with forced air. These lamps run extremely hot, and a small distilled water tank and a heat-absorbing filter were necessary to remove most of the infrared from the beam. The target was illuminated through a ground glass and a gelatin filter that provided more or less monochromatic light at 0.5 μm (5000 Å).

Figure 4 shows visually determined resolution as a function of the focal length at which the pinhole operated—that is, the focal length as defined above by the conjugates p and q. The vertical axis displays normalized resolution limit R, and the horizontal axis, focal length ϕ in units of s^2/λ. The two solid lines represent the theoretical values $R = 1.5$ and $R = 0.61\phi$ for the limiting cases of high and low aperture, respectively.

The solid circles show the experimental data, which were obtained with three different pinholes (diameters 160 μm, 420 μm, and 760 μm) and various values of p and q. As predicted, best resolution is obtained when the focal length of the camera is about equal to s^2/λ, although excursions of 10 or 20% from this value are not excessive. The focusing effect as the camera focal length f approaches s^2/λ is clearly reflected in the data and shows that the pinhole performs best when it is a zone plate encompassing a single Fresnel zone. When the focal length is about $0.7s^2/\lambda$ and the pinhole includes two Fresnel zones, a dark spot appears in the center of the point image; it is curious that this effect does not seriously degrade resolution. (Both Martin

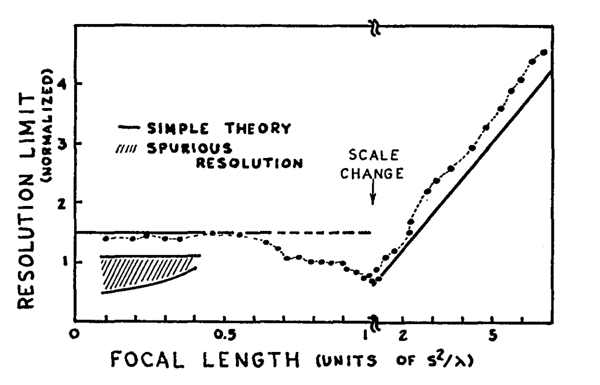


Fig. 4. Resolution limit in units of g vs focal length in units of s^2/λ. Resolution limit is smaller than geometrical shadow when $f = s^2/\lambda$ and pinhole occupies single Fresnel zone. (Because of the definition of f, pinhole radius increases to the left.)

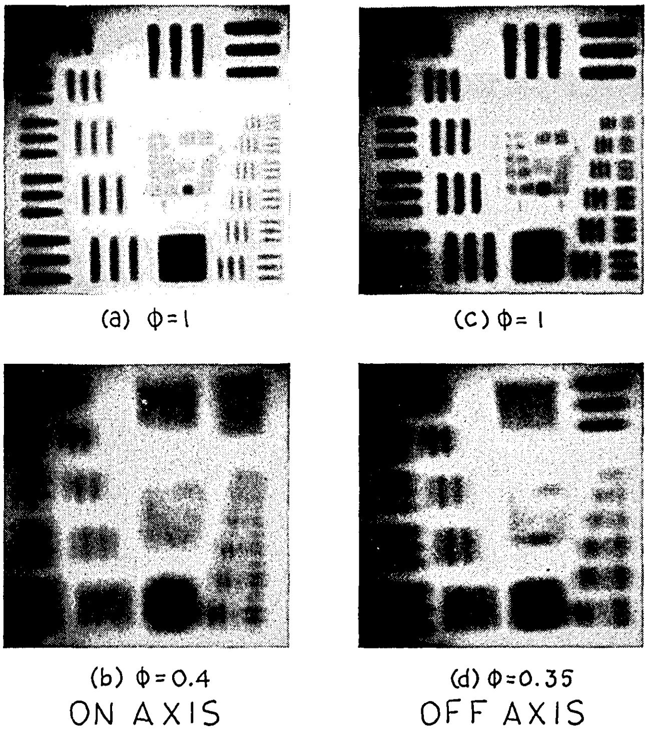


Fig. 5. Images of bar target taken at unit magnification. Upper pictures have optimum pinhole size; lower, two to three times larger. Note spurious resolution and astigmatism.

and Selwyn show photographs of point images for such apertures.)

Figure 5(a) is a photograph of the test pattern exposed with the optimum pinhole size ($\phi = 1$) at unit magnification. The largest bars have a spacing of 1.00 mm.

If the pinhole is expanded to gain exposure, ϕ is reduced. When ϕ is below 0.5, resolution is that predicted by ray optics. At approximately this point, spurious resolution becomes evident, as shown in Fig. 5(b), in which ϕ is about 0.4. The sets of three bars are unresolved, but many appear as sets of two bars. Spurious resolution can be explained by geometrical optics,[14] and these results suggest that this type of spurious resolution is purely a geometrical effect which arises from overlapping of adjacent unresolved bar images.

Depth of Field

The pinhole camera's remarkable depth of field can be appreciated by considering a camera focused at infinity with the pinhole chosen for best resolution; that is, $\phi = 1$ or $q = s^2/\lambda$ (refer to Fig. 2). According to Fig. 4, the image resolution limit is about $0.7s$, since $s = g$ in this case. For nearby objects, the camera behaves as if $\phi < 1$. Figure 4 shows that the resolution limit never exceeds $1.5g$ in this region. For objects distance a q away, g will be $2s$, so the resolution limit in the image will be about $3s$, only four times the value of $0.7s$. But the object, being nearby, is magnified more than are distant objects. This reduces the effect of degrading image resolution and allows us to speak of very large depth of field.

Things are more complicated when the pinhole is optimized for small p, but it is clear from the previous argument that depth of field is considerable.

Astigmatism

I took a second set of data with the pinhole rotated about a vertical axis, so that the target was located 45° off the pinhole axis. (Because the geometrical shadow is used as a normalizing factor, it was not necessary to make the observation plane parallel with the plane of the pinhole.) A one-dimensional theory suggests that the resolution of horizontal bars would be unchanged, and that of vertical bars found by replacing s with $s \cos 45° = 0.7s$. The dashed line in Fig. 6 shows the resolution predicted accordingly for vertical bars, plotted on the same scale as that of Fig. 4. Best resolution for vertical bars is anticipated when $f = (s \cos 45°)^2/\lambda$ or when $\phi = 0.5$. The solid line, for horizontal bars, is taken directly from Fig. 4.

Data points are shown as open circles for horizontal bars and crosses for vertical bars. Agreement with the theory is poor for $\phi > 1$, where the detailed shape of the aperture is important. What is interesting, however, is that there is a region around $\phi = 0.8$ where astigmatism is absent. This corresponds to a pinhole about 10% larger than the optimum for on-axis resolution. Astigmatism is not too serious when $\phi = 1.0$, as is also clear from Fig. 5(c). Figure 5(d) was taken with a larger pinhole, $\phi = 0.35$. Astigmatism is very obvious and in fact similar sets of bars show resolution in one direction and spurious resolution in the other.

Finally, I combined a 35-mm camera body, extension tubes, and black tape with a pinhole to make a pinhole camera. The focal length is 90 mm and ϕ is about 1 at 5000 Å. The pinhole is accordingly 420 μm in diameter, for an $f/200$ system. With a fast film such as Kodak Tri-X, it is possible to hand-hold the camera on a cloudy

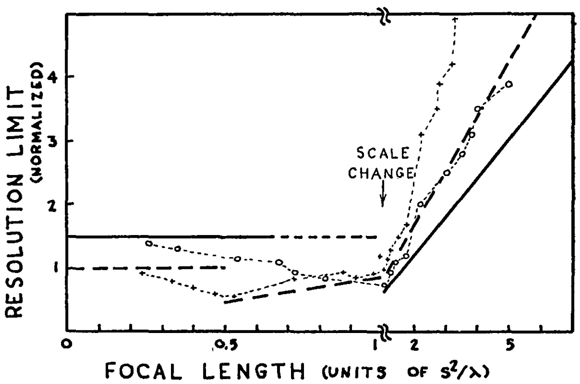


Fig. 6. Normalized resolution limit as in Fig. 4. Target is horizontally off axis. Crosses represent vertical bars; open circles, horizontal bars. Astigmatism is absent for a small range of normalized focal lengths.

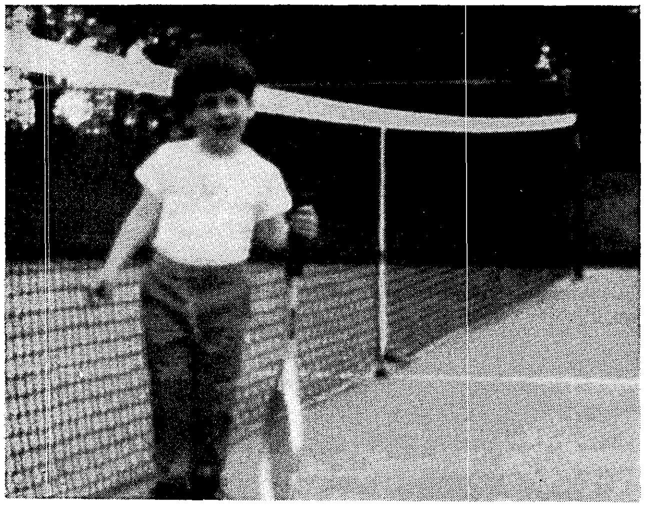


Fig. 7. Snapshot taken with pinhole camera to show soft focus and large depth of field.

day. Figure 7 reproduces a snapshot of a continuous tone object taken with this camera, and shows the kind of soft focus and considerable depth of field that are inherent in the pinhole camera.

Conclusion

I have noted that the pinhole camera can usefully be combined with modern light sources and detectors in a number of relatively specialized applications. (If a laser source is used, we can anticipate that image quality will be worsened by a factor of 2 or more.[15]) To help examine the capability of the instrument under different conditions, I have introduced the focal length of the camera, defined by the object and image distances. The camera gives best resolution when this focal length is equal to the focal length s^2/λ of the pinhole itself.

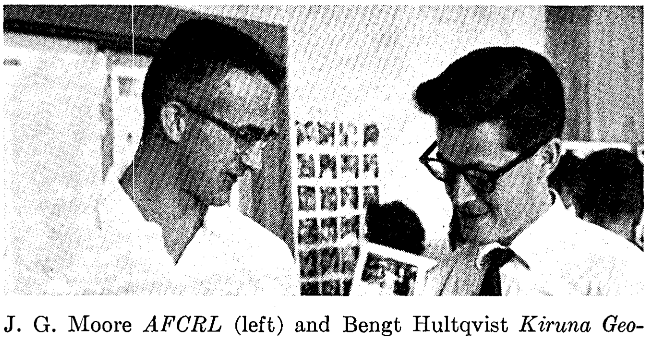
Because the resolution limit can be made at least as small as the pinhole, it is theoretically possible to achieve any on-axis image resolution. This can be done by first choosing a pinhole sufficiently small and then fixing the focal length of the camera at s^2/λ. If the pinhole is small, then the focal length will be short and the image itself small. Object resolution will still be low, since it is, after all, limited by diffraction from a pinhole with a high f/number.

Choice of pinhole diameter fixes both focal length and format, and a compromise will likely be necessary. Often the best pinhole will be the one with the highest aperture that will give the required resolution. If spurious resolution must be avoided, ϕ must exceed 0.5.

In wide-angle applications, the same criteria hold, except where astimatism cannot be tolerated. In that case, Fig. 6 suggests that ϕ be chosen somewhat above the mean of $\cos^2 a$ and 1.0 to get good focusing in both meridians.

References

1. E. Mach, *The Principles of Physical Optics* (Dover, New York, orig. publ. 1926).
2. P. A. Newman and V. E. Rible, Appl. Opt. **5**, 1225 (1966).
3. J. M. Fjeld, Jr., J. Soc. Motion Pict. Telev. Eng. **74**, 320 (1965).
4. A. H. Gallas, C. A. Gilbert, and A. B. Hitterdal, J. Soc. Motion Pict. Telev. Eng. **74**, 321 (1965).
5. M. J. Bernstein and F. Hai, Rev. Sci. Instrum. **41**, 1843 (1970), and references therein.
6. J. H. Waddell, Res./Development **14**, 26 (1963).
7. E. W. H. Selwyn, Photogr. J. **90B**, 47 (1950).
8. L. C. Martin, *Technical Optics* (Pitman, New York, 1959), Vol. 1, pp. 90–91.
9. A. C. Hardy and F. Perrin, *The Principles of Optics* (McGraw-Hill, New York, 1932), pp. 124–126.
10. R. Kingslake, *Lenses in Photography*, rev. ed. (A. S. Barnes and Co., New York, 1963), pp. 60–62.
11. G. O. Reynolds and J. H. Ward, Soc. Photo-Optical Instrum. Eng. J. **5**, 3 (1966); K. Sayanagi, J. Opt. Soc. Am. **57**, 1091 (1967).
12. R. E. Swing and D. P. Rooney, J. Opt. Soc. Am. **58**, 629 (1968).
13. R. S. Longhurst, *Geometrical and Physical Optics*, 2nd ed. (Wiley, New York, 1967).
14. E. L. O'Neill, *Introduction to Statistical Optics* (Addison-Wesley, Reading, Mass., 1963), pp. 25–27.
15. M. Young, B. Faulkner, and J. Cole, J. Opt. Soc. Am. **60**, 137 (1970).



J. G. Moore *AFCRL* (left) and Bengt Hultqvist *Kiruna Geophysical Observatory*, photographed by Sam Silverman *AFCRL* during the September 1961 International Conference on Cosmic Rays and the Earth Storm. Dr. Moore was at NOTS in those days.