Depth Recovery from Blurred Edges

Muralidhara Subbarao
Natarajan Gurumoorthy

Department of Electrical Engineering
State University of New York at Stony Brook
Stony Brook, NY 11794

Abstract

The image formed by a camera having a low depth-of-field contains information about depth of objects in the scene. Objects at a particular distance are focused whereas other objects are blurred by different degrees depending on their distance. A new method is described for recovering depth from a measure of the degree of blur of an edge. Closed-form solution is derived for an edge which is a step discontinuity of intensity in the focused image. The method is simpler and more general than two earlier methods proposed by Pentland (1987) and Grossman (1987). The method has been verified experimentally. It is found to provide useful depth information.

1. Introduction

Finding the distance of objects in a scene from visual information is an important problem in machine vision. There are several methods by which this can be achieved, stereo vision being one example. Recently a new approach based on the depth information inherent in an image formed by a camera has drawn the attention of some researchers (Pentland, 1982, '85, '87; Grossman, 1987; Subbarao, 1987a-c). For a camera with a lens of focal length f the relation between the position of a point in the scene and the position of its focused image is given by the well known lens formula

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v}, \qquad (1.1)$$

where u is the distance of the object, and v is the distance of the image (see Figure 1). Therefore, in the image formed by a camera, only objects at a certain distance are in focus; other objects are blurred by varying degrees depending on their distance. The distance of an object that is in focus can be recovered easily using the lens formula (e.g.: Horn, 1968; Jarvis, 1983; Schlag *et al,* 1983; Krotkov, 1986). Here we consider the more general problem of *finding the distance of an object that may or may not be in focus.*

The basic framework of our approach is as follows. The observed image of an object is modeled as the result of convolving the focused image of the object with a point spread function. This point spread function depends both on the camera parameters and the distance of the object from the camera. The point spread function is taken to be rotationally symmetric (this is largely true of cameras with circular apertures). The line spread function corresponding to this point spread function is computed from a blurred step-edge. A measure of the *spread* of the line spread function is estimated from its second central moment. This spread is shown to be related linearly to inverse distance. The constants of this linear relation are determined through a simple camera calibration procedure. Having computed the spread, distance of the object is obtained from the linear relation.

2. Relation to previous work

Pentland (1982, '85, '87) and Grossman (1987) both addressed the problem of recovering depth from blurred edges. Pentland modeled a blurred edge as the result of convolving a focused image with a *Gaussian point spread function.* He showed that if $C(x,y)$ is the laplacian of the observed image then the spread σ of the Gaussian is related to $C(x,y)$ by

$$\ln\left[\frac{b}{\sqrt{2\pi\sigma^3}}\right] - \frac{x^2}{2\sigma^2} = \ln\left|\frac{C(x,y)}{x}\right|, \qquad (2.1)$$

where b is the magnitude (or "height") of the step edge and x,y the image coordinate system with its origin on the edge and x-axis perpendicular to the edge. In the above equation b and σ are the only unknowns. He solved for these unknowns by formulating the above equation as a linear regression in x^2. The depth was then computed from σ. Pentland applied his method to an image of a natural scene and showed that depth of edges could be classified as being small, medium, or large.

Our approach differs from Pentland's in two respects. First, as we shall see later, the computational method is simpler. The solution for b and σ are given in closed-form. Second, and more important, our method assumes only that the point spread function of the camera be rotationally symmetric. It is not restricted to the case where the point spread function is a Gaussian.

Grossman (1987) showed experimentally that useful depth information can be obtained from blurred edges. He considers, in addition to step-edges, edges of finite width (e.g.: ramp edge). Although Grossman

does not provide a theoretical justification for his computational algorithm, he did have the right intuitions. Almost all the major assumptions of his method have been shown to be correct by Pentland's work and our work. Only one of his stated assumption that "It is possible, in most situations, to create an image with a useful range of blur." is not required. By this assumption Grossman appears to imply that his method requires the presence of several edges in the image with different amounts of blur caused by the objects being at different distances. This requirement can be eliminated by using an appropriately calibrated camera.

3. Basic theory

3.1 Point spread function of a lens

Let P be a point on a visible surface in the scene and p be its focused image (see Figure 1). The relation between the positions of P and p is given by the lens formula (1.1). If P is not in focus then it gives rise to a circular image called the *blur circle* on the image detector plane. From simple plane geometry (see Figure 1) and the lens formula we can show that the diameter d of the blur circle is given by

$$d = D\ s\ \left[\frac{1}{f} - \frac{1}{u} - \frac{1}{s}\right] \tag{3.1}$$

where D is the diameter of the lens and s is the distance from the lens to the image detector plane. Note that d can be either positive or negative depending on whether $s \geq v$ or $s < v$. In the former case the image detector plane is *behind* the focused image of P and in the latter case it is *in front* of the focused image of P. According to geometric optics, the intensity within the blur circle is approximately constant given by

$$h_1(x,y) = \begin{cases} \dfrac{4}{\pi d^2} & \text{if } x^2+y^2 \leq \dfrac{d^2}{4} \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

But due to diffraction and non-idealities of lenses (Horn, 1986; Schreiber, 1986; Pentland, 1987; Subbarao, 1987a-c) an alternative model is suggested for the intensity distribution given by a two-dimensional Gaussian

$$h_2(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2}\frac{x^2+y^2}{\sigma^2}} \tag{3.3}$$

where σ is the spread parameter such that

$$\sigma = k\ d \quad \text{for } 0 < k. \tag{3.4}$$

(k is a constant of proportionality characteristic of a given camera; it is determined initially by an appropriate calibration procedure.) Now the blurred image of point P is actually the *point spread function* of the camera system. Therefore, for a linear shift-invariant camera system (cf. Rosenfeld and Kak, 1982), an observed image is the result of convolving the focused image with the camera's point spread function. (The *focused image* of a scene for a given position of the image detector plane can be defined as follows. For any point p on the image detector (see Figure 1), consider a line through that point and the second principal point. Let P be the point on a visible surface in the scene whose focused image lies on this line. Let p' be the focused image of P. Then the intensity at p is the intensity of the focused image at p'.)

$h_2(x,y)$ in equation (3.3) is defined in terms of σ and therefore is different for points at different distances from the lens plane. The "volume" of both $h_1(x,y)$ and $h_2(x,y)$ can be shown to be unity. (The volume of the point spread function of a non-absorbing lens is unity irrespective of the form of the point spread function.)

3.2 Recovering the line spread function of a camera

Let $f(x,y)$ be a step edge along the y-axis on the image plane. Let a be the image intensity to the left of the y-axis and b be the height of the step. The image can be expressed as

$$f(x,y) = a + b\ u(x) \tag{3.5}$$

where $u(x)$ is the standard *unit step function*. If $g(x,y)$ is the observed image and $h(x,y)$ is the camera's point spread function then we have

$$g(x,y) = h(x,y) * f(x,y)\ , \tag{3.6}$$

where * represents the convolution operation. As mentioned above, the "volume" of the point spread function of a camera is unity, i.e.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\ dx\ dy = 1\ . \tag{3.7}$$

Now consider the derivative of g along the gradient direction. Since differentiation and convolution are linear operations they commute. Therefore we have

$$\frac{\partial g}{\partial x} = h(x,y) * \frac{\partial f}{\partial x} \tag{3.8}$$

$$= h(x,y) * b\,\delta(x)$$

where $\delta(x)$ is the *dirac delta function* along the x axis. Therefore we have

$$\frac{\partial g(x,y)}{\partial x} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\xi,\eta)\ b\,\delta(x-\xi)\ d\xi d\eta \tag{3.9}$$

$$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} h(\xi,\eta)\ b\,\delta(x-\xi)\ d\xi\right] d\eta$$

$$= b\int_{-\infty}^{\infty} h(x,\eta)\ d\eta$$

The last term on the right hand side of the above expression is a *line integral* along a line parallel to the y-axis. Let $\theta(x)$ denote this line integral, i.e.

$$\theta(x) = \int_{-\infty}^{\infty} h(x,y)\, dy. \tag{3.10}$$

$\theta(x)$ is in fact the *line spread function* of the camera system. Thus we have

$$\frac{\partial g(x,y)}{\partial x} = b\ \theta(x). \tag{3.11}$$

Now consider the integral

$$\int_{-\infty}^{\infty} \frac{\partial g(x,y)}{\partial x}\, dx = \int_{-\infty}^{\infty} b\ \theta(x)\, dx \tag{3.12}$$

$$= b \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\, dy\, dx$$

$$= b \qquad \text{(from 3.7)}.$$

Therefore given the observed image $g(x,y)$, we can obtain the line spread function $\theta(x)$ from the expression

$$\theta(x) = \frac{\frac{\partial g}{\partial x}}{\int_{-\infty}^{\infty} \frac{\partial g}{\partial x}\, dx}. \tag{3.13}$$

If the shape of the camera aperture is circular then it is reasonable to assume that the point spread function is circularly symmetric. In this case the line spread function $\theta(x)$ is identical for all orientations of the image coordinate system. In the following discussion we assume that *the point spread function is circularly symmetric.*

3.3 Relation between the spread of the line spread function and depth

We shall define the *spread parameter* of a line spread function to be *the square root of the second central moment of the line spread distribution.* (This corresponds to the "standard deviation" of the distribution of the line spread function. It can also be interpreted as the "radius of gyration" about the "center of mass".) Therefore, if σ_l represents the spread parameter, then we have

$$\sigma_l^2 = \int_{-\infty}^{\infty} (x-\bar{x})^2\, \theta(x)\, dx \tag{3.14}$$

where $\bar{x}$ is the "center of mass" of the distribution defined by

$$\bar{x} = \int_{-\infty}^{\infty} x\ \theta(x)\, dx. \tag{3.15}$$

(Note that the "total mass" of the distribution is unity, i.e.

$$\int_{-\infty}^{\infty} \theta(x)\, dx = 1 \tag{3.16}$$

).

Now we consider the relation between σ_l and the distance u. Suppose that the camera is ideal in the sense of geometric optics. Then the point spread function $h_1(x,y)$ (see equation 3.2) characterizes the camera exactly. The line spread function (obtained from equation 3.10) corresponding to h_1 is

$$\theta_1(x) = \begin{cases} \frac{4}{\pi d^2}\sqrt{d^2-4x^2} & \text{if } |x| \leq \frac{d}{2} \\ 0 & \text{otherwise.} \end{cases} \tag{3.17}$$

The spread parameter of this line spread function can be shown to be $d/4$. Similarly we find that the line spread function corresponding to h_2 in equation (3.3,3.4) is

$$\theta_2(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{x^2}{\sigma^2}}. \tag{3.18}$$

The spread of this line spread function is σ itself. These results suggest the following relation between σ_l and d:

$$\sigma_l = k_l d \tag{3.19}$$

for some constant k_l. Experimental results indicate that this relation holds quite well for the camera used in our experiments. From equations (3.1,3.19) we have

$$\sigma_l = k_l D\ s \left[\frac{1}{f} - \frac{1}{u} - \frac{1}{s}\right]. \tag{3.20}$$

This equation suggests that, for a given camera parameter values, the relation between σ_l and the distance u can be expressed in the form

$$\sigma_l = mu^{-1} + c \tag{3.21}$$

where m, c are some camera constants. These constants can be determined initially once and for all by a suitable camera calibration method. The important point to note here is that *the spread parameter is linearly related to inverse distance.* Therefore, having determined the spread parameter from the observed image, the distance can be easily computed.

In a polar coordinate system (r,ϕ), the spread parameters of a point spread function and its corresponding line spread function can be defined as

$$\sigma^2 = \int_0^{\infty}\int_0^{2\pi} h(r)\, r^3\, d\phi\, dr \tag{3.22a}$$

and

$$\sigma_l^2 = \int_0^{\infty}\int_0^{2\pi} h(r)\, r^3 \cos^2\phi\, d\phi\, dr \tag{3.22b}$$

respectively. It can be easily shown that

$$\sigma = \sqrt{2}\sigma_l. \tag{3.23}$$

The above relation in conjunction with the experimental verification of equation (3.21) justifies the assumption $\sigma = k\ d$ in equation (3.4)

4. Experiments and computational steps

The goal of our experiments was two fold: (i) to

verify the applicability of our mathematical model to practical camera systems, and (ii) to test the usefulness of the method in practical applications.

Black and white papers were pasted on a cardboard to create a step discontinuity in reflectance along a straight line. Nine 64×64 images of the cardboard were acquired with a CCD camera (Four of the pictures are shown in Figure 2). The camera was a Panasonic WV-CD50, focal length 16 mm, aperture diameter 11.4 mm, and pixel size 0.013 mm × 0.017 mm. The distance of the cardboard from the camera was different for different images but the camera parameters were the same for all images. The camera was focused to 24", and the images were acquired for distances of 8", 11", 14", 17", 20", 24", 33", 48" and 100".

For each image, the first derivative was computed along the intensity gradient by simply taking the difference of gray values of adjacent pixels. In our experiments the gradient was along rows of the image matrix (i.e. the gray values of pixels in a column were approximately constant). The resulting image was then cut into 4 strips of equal size along the rows. In each row, the approximate location of the edge was computed by finding the first moment of the distribution of the image derivative, i.e. if $\bar{i}$ is the edge location and $g_x(i)$ is the image derivative, then

$$\bar{i} = \frac{\sum_{i=0}^{64} i\, g_x(i)}{\sum_{i=0}^{64} g_x(i)}. \qquad (4.1)$$

The following step was included to reduce the effects of noise. Each row was traversed on either side of position $\bar{i}$ until a pixel was reached where either $g_x(i)$ was zero or its sign changed. All pixels between this pixel (where, for the first time, g_x became zero or its sign changed) and the pixel at the row's end were set to zero. From the resulting $g_x(i)$, the line spread function was computed using

$$\theta(i) = \frac{g_x(i)}{\sum_{i=0}^{64} g_x(i)}. \qquad (4.2)$$

The location of the edge $\bar{i}$ was then recomputed using equation (4.1). The spread σ_l of the line spread function was computed from

$$\sigma_l = \pm \sqrt{\sum_{i=0}^{64} (i-\bar{i})^2 \theta(i)}. \qquad (4.3)$$

The computed values of σ_l for adjacent rows were found to differ by only about two percent. The average $\bar{\sigma}_l$ was computed in each image strip (16 pixels wide). The resulting values are plotted against inverse distance in Figure 3.

In Figure 3, the values of $\bar{\sigma}_l$ for the four strips of each image almost exactly coincide and therefore are not distinguishable. This implies that the line spread function of our camera is shift-invariant. The image corresponding to $u=24$ inches was visually judged to be the sharpest possible image obtainable from our camera. The experiments are in agreement with our visual judgement in that σ_l is a minimum for this image. Let σ_{omin} denote this minimum observed σ_l. On either side of this minimum the plot is very nearly linear, and further the slopes of the lines on either side are almost exactly equal. This verifies equation (3.21) which predicts a linear relation between σ_l and u^{-1}. Another interesting observation is that although the object in the scene had a perfect step edge, its sharpest image had a finite width of about two pixels as indicated by σ_{omin} being one pixel. Let $\sigma_l=\sigma_{amin}$ at the point where the straight lines on either side of the minimum intersect. In Figure 3 we find that $\sigma_{amin} \simeq 0.5$ pixels and that the intersection point is almost vertically below the actual observed minimum. We interpret σ_{omin} and σ_{amin} as follows. We attribute σ_{amin} to the inherent blur in the image formed on the CCD array the blur being caused by lens aberrations. The difference $\sigma_{omin}-\sigma_{amin}$ is introduced by discrete spatial and gray level sampling of the image.

Equation (4.3) suggests a two fold ambiguity for σ_l. This ambiguity can be overcome by setting the distance between the lens and the image detector plane to be equal to the focal length, i.e. $s=f$. In this case σ_l is always negative (see equation 3.20). Therefore a unique solution is obtained.

We found that noise cleaning is important in our experiments. A small non-zero value of image derivative caused by noise at pixels far away from the position of the edge affects the estimation of σ_l considerably. The farther a pixel is from the edge, the higher its effect on the final result (see equation 4.3) because the distance is squared.

The effective range of our method depends on the constants m, c in equation (3.21) and the image quality in terms of spatial and gray level resolution. This method is more effective for objects at shorter distances than at longer distances because blur varies linearly with inverse distance. The maximum distance of the cardboard in our experiments was about 8 feet. The range can be increased by using a custom designed camera which produces high quality images.

5. Conclusion

We have described a simple method for depth recovery by measuring the degree of blur of a step-edge. The depth recovery is accurate for nearby objects and approximate for far away objects.

One limitation of our method is that it is restricted to isolated step edges; presence of other edges nearby (within a distance of about twice the spread parameter of the line spread function) affects depth estimation. It is possible to extend our approach to patterns of focused images other than the step edge (e.g. a ramp edge). But

the mathematics and the computational methods appear to become complicated.

Acknowledgement: We thank Dr. Bill Sakoda for his help in acquiring image data for experiments.

6. References

Grossman, P. 1987 (Jan.). Depth from focus. *Pattern Recognition Letters 5.* pp. 63-69.

Horn, B.K.P. 1968. Focusing. Artificial Intelligence Memo No. 160, MIT.

Horn, B.K.P. 1986. *Robot Vision.* McGraw-Hill Book Company.

Jarvis, R. A. March 1983. A perspective on range finding techniques for computer vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* PAMI-5, No. 2, pp. 122-139.

Krotkov, E. 1986. Focusing. MS-CIS-86-22. Grasp lab 63. Dept. of Computer and Information Science. University of Pennsylvania.

Ligthart, G. and F. C. A. Groen. 1982. A comparison of different autofocus algorithms. *Proceedings of the International Conference on Pattern Recognition.*

Ogle, K.N. 1968. *Optics.* Chapter VII, Section 2. Charles C Thomas . Publisher. Springfield, Illinois.

Pentland, A.P. 1982. Depth of scene from depth of field. *Proceedings of DARPA Image Understanding Workshop.* Palo Alto.

Pentland, A. P. 1985. A new sense for depth of field. *Proceedings of International Joint Conference on Artificial Intelligence,* pp. 988-994.

Pentland, A.P. 1987. A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. PAMI-9, No. 4, pp. 523-531.

Rosenfeld, A., and A.C. Kak. 1982. *Digital Picture Processing,* Vol. I. Academic Press, Inc.

Schlag, J.F., A.C. Sanderson, C.P. Neuman, and F.C. Wimberly. 1983. Implementation of automatic focusing algorithms for a computer vision system with camera control. CMU-RI-TR-83-14. Robotics Institute. Carnegie-Mellon University.

Schreiber, W.F. 1986. *Fundamentals of Electronic Imaging Systems.* Springer-Verlag. Section 2.5.2.

Subbarao, M. 1987a Direct recovery of depth-map: Differential Methods. *Proceedings of IEEE computer society workshop on computer vision, Miami Beach, Dec. 1987,* pp 58-65.

Subbarao, M. 1987b (April). Direct Recovery of Depth-map II: A New Robust Approach. Technical report 87-03. Image Analysis and Graphics Laboratory, SUNY at Stony Brook.

Subbarao, M. 1987c (May). Progress in research on direct recovery of depth and motion. Technical report 87-04. Image Analysis and Graphics Laboratory, SUNY at Stony Brook.

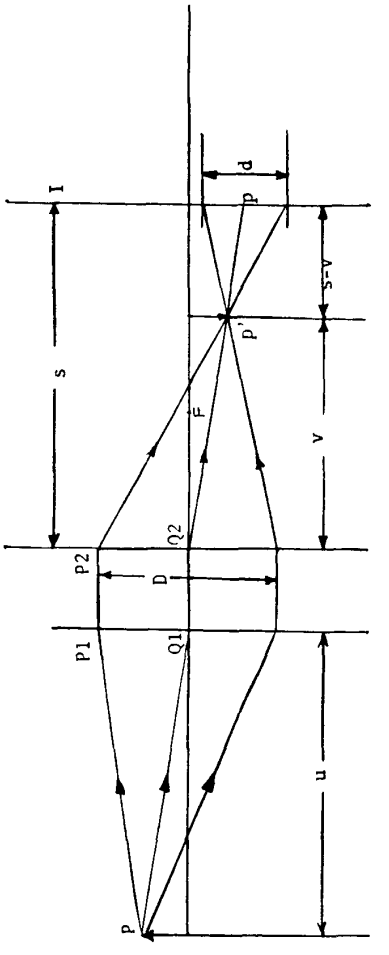
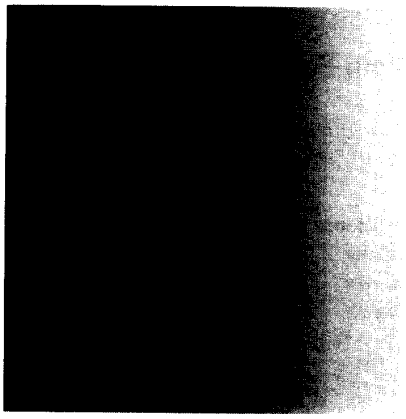


Figure 1. The camera geometry and camera parameters.
P1: first principal plane P2: second principal plane
P : object point p: image point I: image detector plane
s,f,D: camera parameters u: object distance v: image distance
d: blur circle diameter Q1,Q2: first & second principal points



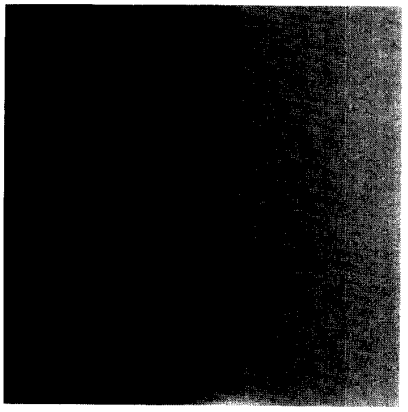
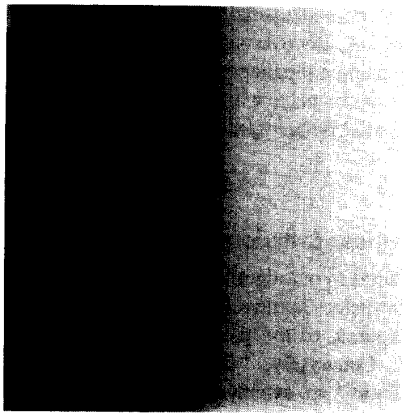
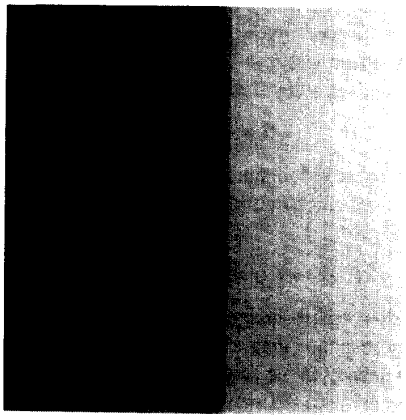


Figure 2





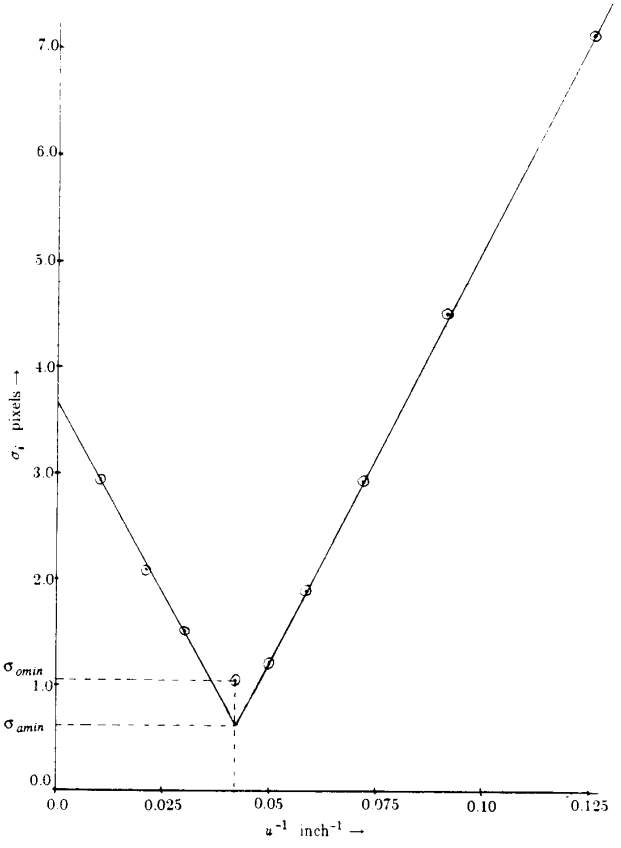


Figure 3