## 3-D Images on a 2-D Display

So far we have used intensity as a third dimension to view how a quantity may vary as a function of two variables, which we have denoted $f(x,y)$. In each case we have plotted the intensity on our display grid assuming each point's location in pixels is directly proportional to its coordinates $x$ and $y$. But this isn't the only way to present these data — and sometimes we want to display 3-D data on our 2-D display. Let's look at some ways to do these things.

In addition, sometimes we want to look at vector functions of 2-D data, and again we will have to display some reduced version of our initial data set. An example of this is topographic data, where we might have two quantities at each $x,y$ location, say the altitude of a point and its color or brightness in an image.

Topography. Let's begin by defining simple topography, which we can define as a 2-D function $f(x,y)$ giving the height above some reference of locations on the ground. For example, a topographic map of the local Stanford area would have almost the same height for all the points in the Main Quad, and a local maximum point for the spot corresponding to the top of the hill where the Dish sits. Here $x$ and $y$ give the location of each point, say its

latitude and longitude. The value of the function might
be meters above sea level or some similar units.

How can we present these data? The method most like
what we have used so far would be to convert each
elevation to a gray level, say for black to correspond
to the lowest height in the image, and white for the
highest. Our local area displayed this way might
look like this:
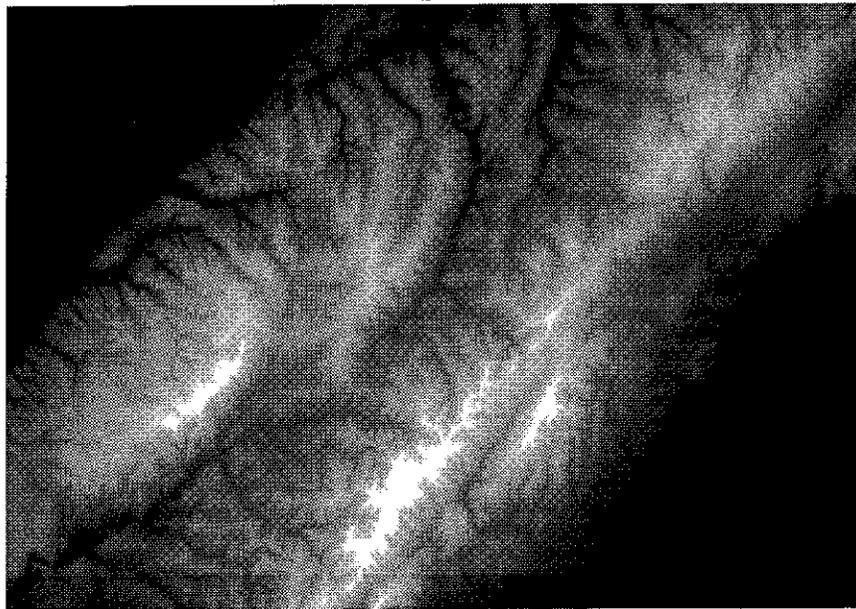


We could also code the heights with color, instead
of gray. The gray image above certainly shows the
high and low areas in a large-scale sense, but it
actually isn't very easy to visualize or interpret
details of the image. It is especially hard to see
fine-scale details.

<u>Shaded relief</u>. Digital mapmakers have been aware of the perceptual limitations of the straight intensity map, and often use the technique of shaded relief to accentuate the details of a topographic image. Although cartographers refer to these display methods as slope maps, we can recognize shaded relief as a mathematical function that accentuates small-scale relief.

Suppose we have a function $f(x,y)$, and want to emphasize fine-scale structure. We want to use a filter that passes fine-scale features but not large-scale features. Since these relate, as we now know, to high and low spatial frequencies, respectively, we want to filter the function with a high-pass filter.

And we have looked at one version of a high-pass filter already — the derivative operator.

Note that we have even looked at an implementation of a derivative operator: applying a convolutional filter with a mask that looks like $[1 \ -1]$. Applying this filter to a discrete image $f(i,j)$ results in a new image $g(i,j)$ such that

$$g(i,j) = f(i,j) - f(i+1, j)$$

Since each pixel represents the difference of two neighboring input pixels, the intensity of the new image is proportional to the slope of the underlying terrain.

Displaying the previous image as a shaded relief map, we have
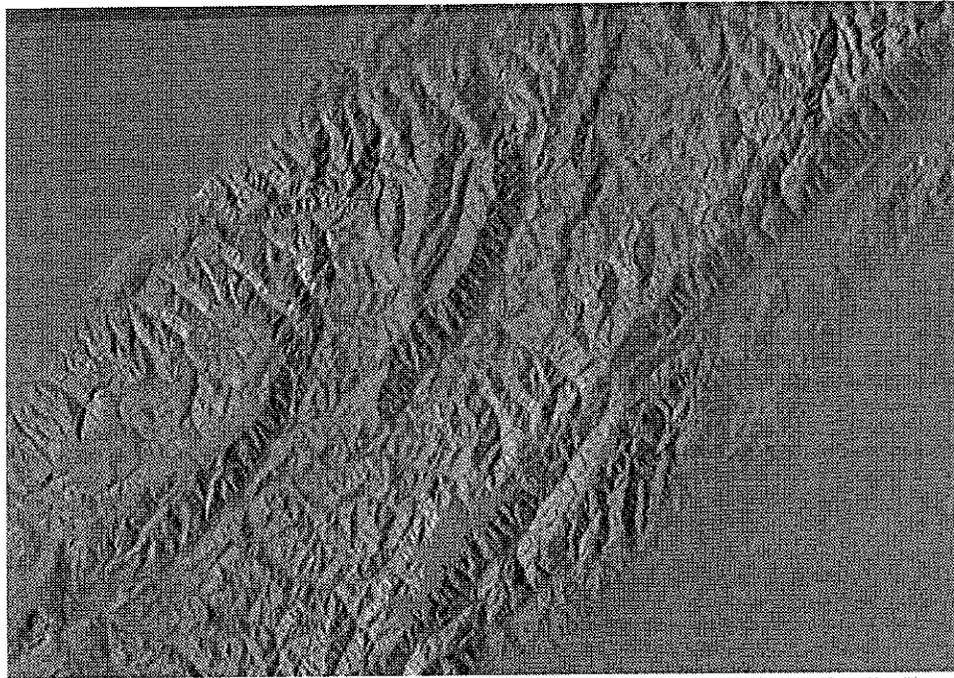


Note how much easier it is to interpret the highs and lows in this image. In fact we have less information in this picture than before, as we now have only the pixel differences. But the use of image contrast is much improved, and our brain can process the data much more accurately.

Note that we could have used many different convolution masks instead of the $[1 \quad -1]$ mask, such as $[-1 \quad 1]$, $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, or even $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Each of

these represents a derivative calculated in a different direction, so that a different component of the slope is displayed.

<u>Practical aspects of shaded relief generation</u>. There are several practical points to keep in mind when generating shaded relief images. For one thing, slopes are both positive and negative, while topography is usually only a positive quantity. To represent both, we usually represent zero slope as a neutral gray level of 128, with values increasing for upslopes and decreasing for downslopes. An equation for a shaded relief map might be

$$g(i,j) = F \cdot \left[ f(i,j) - f(i-1,j) \right] + 128$$

where the constant $F$ is chosen to make the range of slopes map comfortably into our 8-bit display.

Another issue is related to one property of the high-pass filter, its propensity to magnify high-frequency signals. Sampling or other measurement noise tends to be broad-band, or contain a lot of high-frequency components. Hence a simple shaded-relief image may appear rather noisy and unaesthetic. We can overcome this limitation if we recognize that our derivative is calculated in one dimension, hence we can integrate in the orthogonal direction to keep the overall noise level about the same. This means using convolution kernels such as

$$\begin{bmatrix} 1 & -1 \\ 1 & -1 \\ 1 & -1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ & 1 & 0 & -1 & 0 \\ & & 0 & -1 & 0 & 0 \end{bmatrix}$$

Our overall equation then becomes

$$g(i,j) = F \cdot f(i,j) * m(i,j) + 128$$

where $m(i,j)$ is our convolution mask.

Contour lines. Another time-tested means of display for topographic data is the contour plot. Here points of equal elevation are connected by lines, or contours. Consider the following array of pixels:

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 2 & 1 & 1 \\ 1 & 1 & 2 & 3 & 2 & 1 & 0 \\ 1 & 2 & 3 & 4 & 3 & 2 & 1 \\ 1 & 1 & 2 & 3 & 2 & 2 & 1 \\ 0 & 1 & 1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Contour lines for the 2 and 3 contours might be

or, simply,
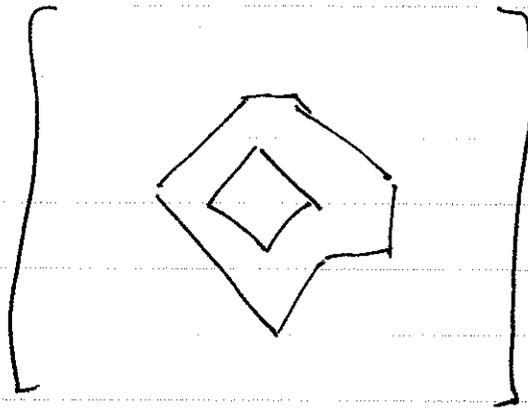


This suggests a mountain peak in the center of the image. The extension to more complicated contours is obvious!



Here it is quite easy to find peaks and valleys. Can you identify the easiest way to hike up the mountains on this island? How many peaks are there? How many saddle points?

Contour maps present a great deal of information with a very simple drawing, and remain a popular way to express elevation data or other 2-D functions.
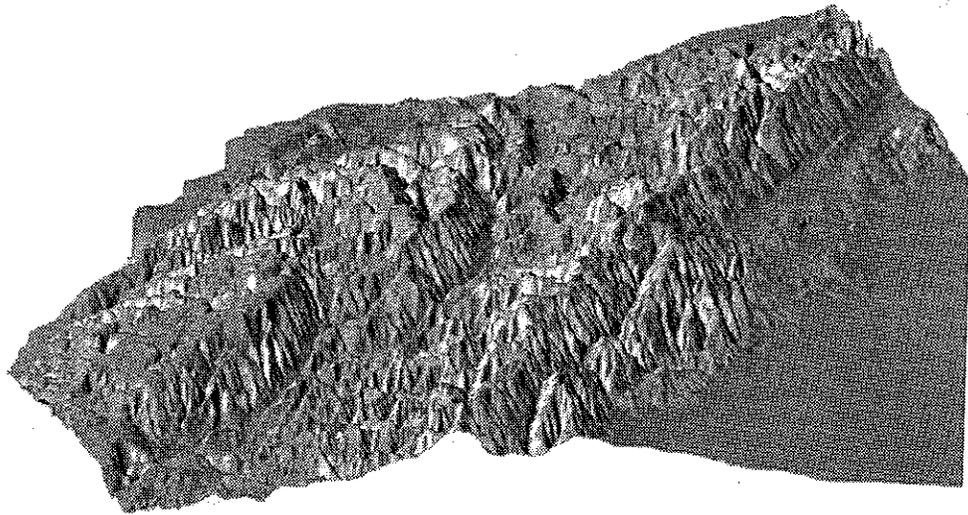
## Projections

    Representing a 3-D object on a 2-D screen or paper requires us to reduce the original object in some respect. In order to illustrate faithfully the desired properties of the original, 3-D distribution, some care must be taken in how to do the projection, because inevitably some information is lost.

    We in fact come across projections in many forms every day. Any photograph represents a projection of one kind, where we use contextual clues such as one

item being placed in front of another or the relative sizes of objects to infer the original 3-D view.

Another common projection is the one we see in conventional x-ray pictures. In an x-ray the integrated tissue density through a body is displayed, and the doctor's knowledge of how the internal organs are arranged provide the additional information needed to interpret the result.

## Perspective projections

Perspective views result from choosing a viewpoint $(x_0, y_0, z_0)$ in a 3-D space, and representing each point in the space on a plane, usually perpendicular to the look direction, where a ray from that point intersects the plane.



Viewpoint
$(x_0, y_0, z_0)$
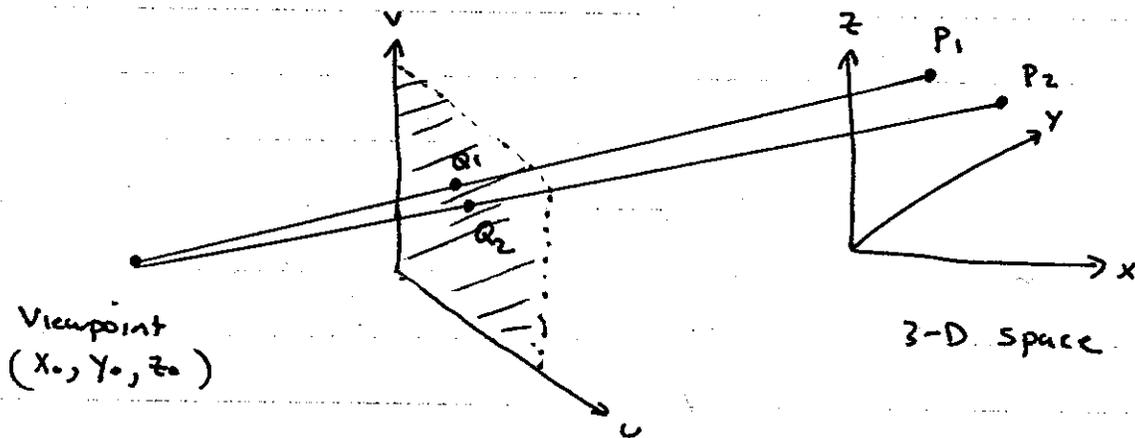
3-D Space

So points $P_1$ and $P_2$ in 3-D space map into points $Q_1$ and $Q_2$ in 2-D space. Varying the viewpoint obviously changes the projection dramatically.

We can derive the mathematical transformation of the coordinates of $P_1$ and $P_2$ to $Q_1$ and $Q_2$ readily,

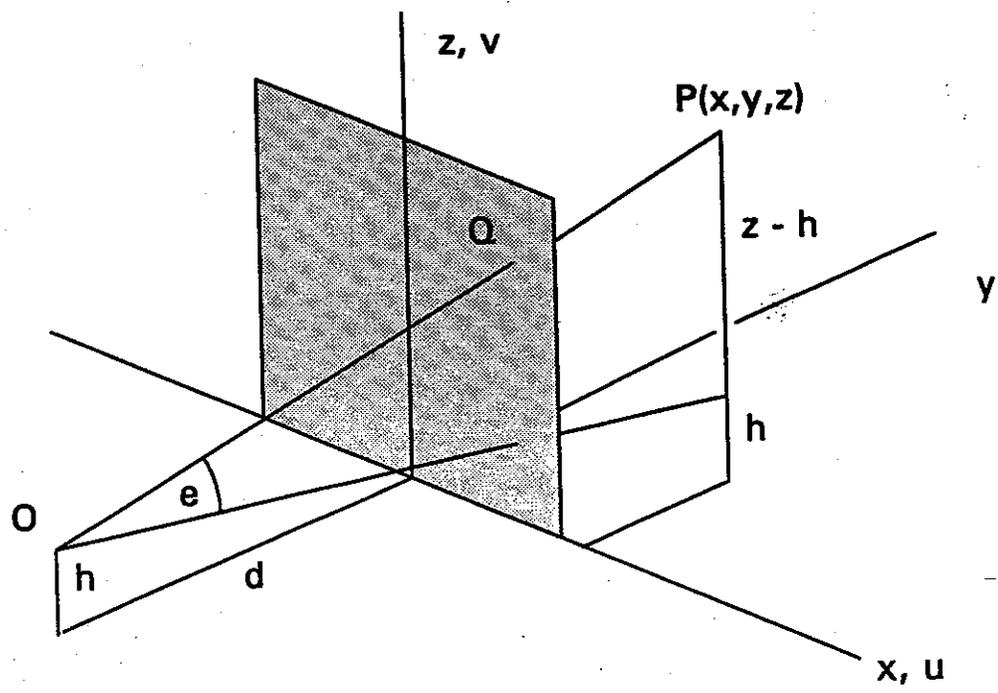especially if we adopt the following set of coordinates :



**Figure 2-18** The relationship between the space coordinates $(x, y, z)$ of the object point $P$ and the picture-plane coordinates $(u, v)$ of the projected point $Q$, when the $(x, z)$-plane is chosen to coincide with the $(u, v)$-plane.

The origin for the eye is located at $(0, -d, h)$; geometrical construction leads to

$$u = \frac{xd}{y+d}$$

$$v = h + \frac{(z-h)d}{y+d}$$

Note that the right hand sides of these equations involve three variables, $x$, $y$, and $z$, while the left hand sides only two, $u$ and $v$. Thus three dimensions are reduced to two.

Is this operation invertible ? What does that mean?

## The pinhole camera

An alternate construction for perspective viewing is derived from a model of a pinhole camera, in which all of the light rays from an object are passed through a single point and then diverge onto an "imaging plane". A little thought shows that this is a special case of the more general projection just described. Here is the geometry:



pinhole

Imaging plane
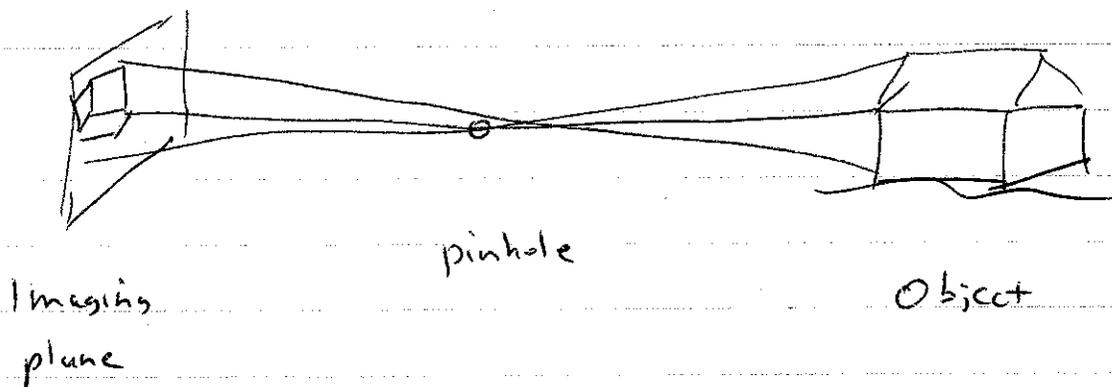
Object

Note that the object appears upside down on the plane, so the final image must be flipped up and down and left and right to be a faithful reproduction.

We can set up a set of coordinates to describe the imaging equations. Let the object be described as a set of points $(x, y, z)$ offset by $y_0$ from a pinhole located at $(0, 0, 0)$. The imaging plane is offset by a distance $d$ in the opposite direction, and is seen to be the plane $y = -d$. Then,

$z$

$(x, y+y_0, z)$

$Y$

$X$

$y_0$

$z'$

$d$

Pinhole at
$(0, 0, 0)$

$x'$   Image point at $(x', -d, z')$
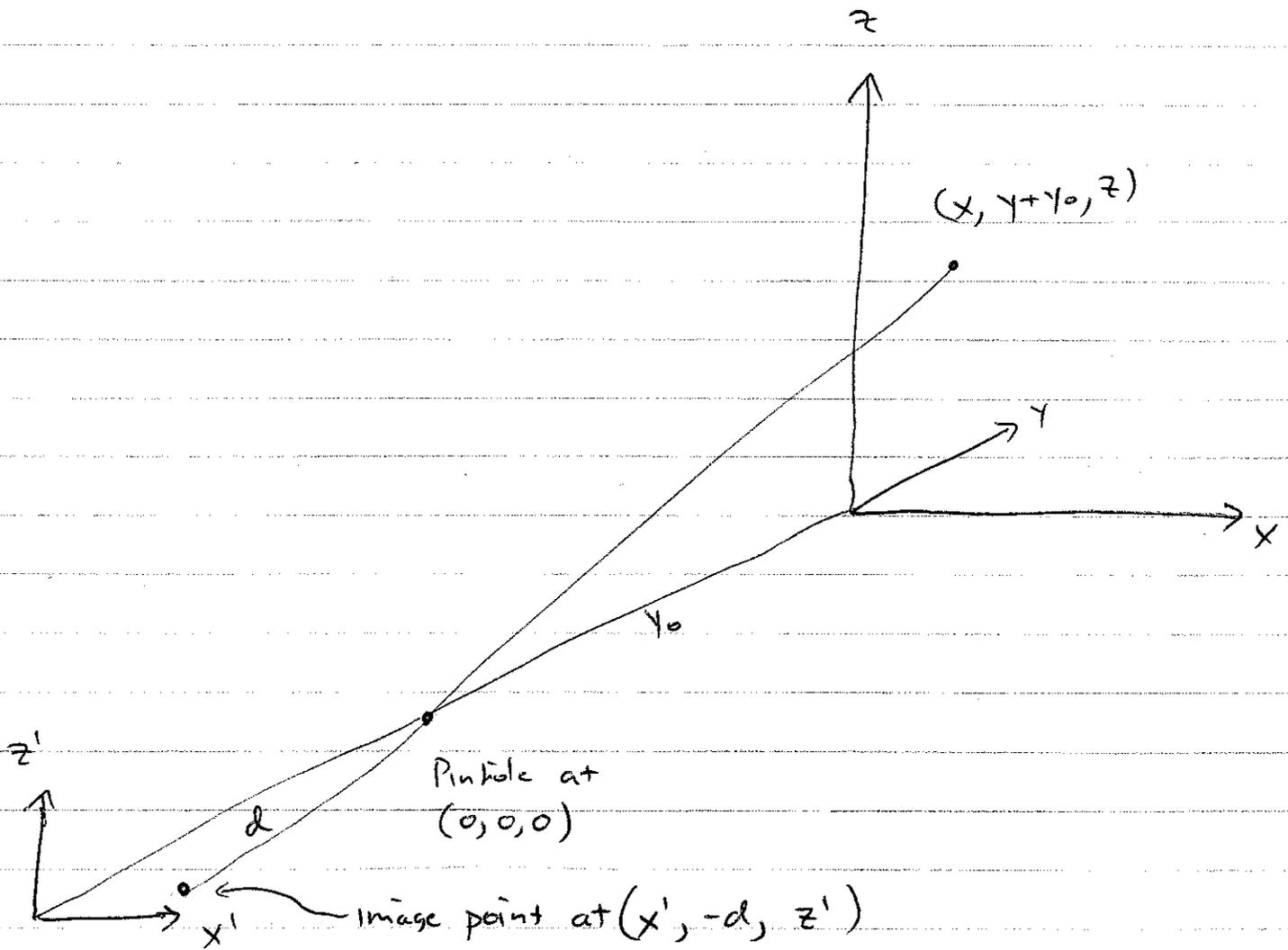
Each point on the object is described by a vector from the pinhole to $(x, y+y_0, z)$. We equate a general form for any point on the line to its location in the $y' = -d$ plane:

$$a \begin{pmatrix} x \\ y+y_0 \\ z \end{pmatrix} = \begin{pmatrix} x' \\ -d \\ z' \end{pmatrix}$$

Simple algebraic solution gives    $a = \dfrac{-d}{y+y_0}$

and hence

$$x' = \frac{-d}{y+y_0} x$$

$$z' = \frac{-d}{y+y_0} z$$

These equations thus map any point $(x, y+y_0, z)$ into the corresponding $(x', z')$, forming the pinhole image.

## Inverse problems with perspective

So it is straight-forward to determine how the 2-D perspective view of an object appears as a function of viewing geometry. Often we are faced with the _inverse problem_, where we are presented with one or more 2-D perspective views and asked to reconstruct the 3-D original object.

We have mentioned previously the use of contextual clues for qualitatively interpreting perspective views. We use many of these in our everyday experiences - name some of them:

      Objects in foreground / background

      Sizes

      Lighting clues

        ⋮

How can some of these be made quantitative?


## Orthographic projections:

One method of aiding quantitative interpretation is orthographic ~~reditions~~ renditions. An _orthographic_ representation is one where the viewpoints are all an infinite distance away, and are all perpendicular to the object plane ~~each other~~. These projections are indispensable for engineering and architectural drawings. Typically the three views used are of the front, side, and top of the item to be displayed.

A significant advantage of the orthographic projection is that all distances are true, that is given a single scale marker all dimensions may be inferred. This doesn't occur for many perspective drawings.
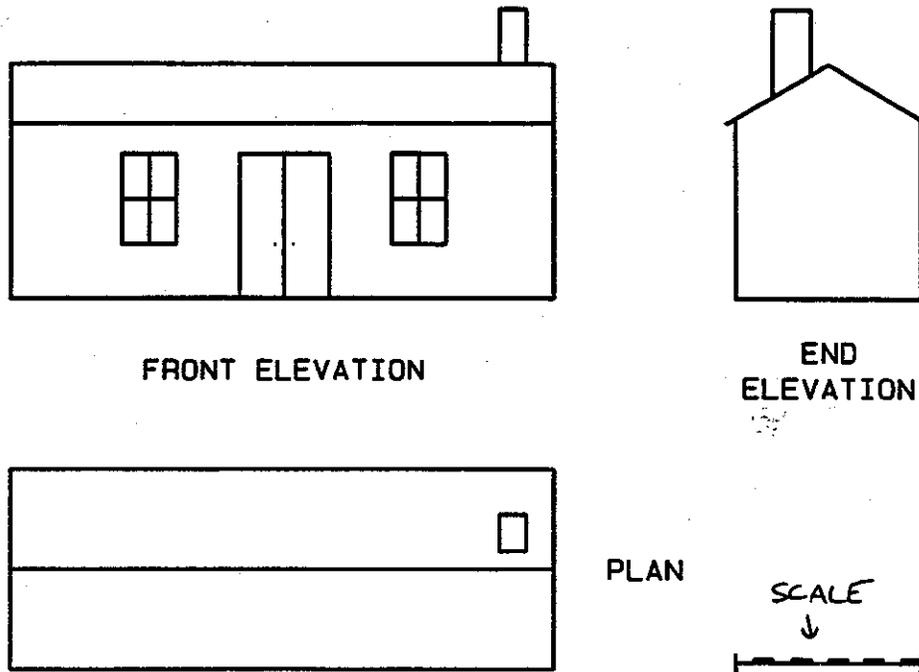
**Figure 2-20**  Orthographic projections in the form of plan, elevation, and side elevation.

## Isometric projection

   A certain orthographic projection, the _isometric_ projection, is taken from a special direction so that the object's principal axes ($x, y,$ and $z$) are all inclined equally. In other words, if $\hat{n}$ represents a vector in the viewing direction;

$$\hat{n} \cdot \hat{x} = \hat{n} \cdot \hat{y} = \hat{n} \cdot \hat{z} \qquad (*)$$

   This turns out to be an angle of $\tan^{-1}\sqrt{2} = 54.7°$ between an axis and the look direction, and the angle between any axis and the picture plane is $35.3°$.

   Because the projections in $(*)$ are all equal, any given distance in the image along one of the axes projects equally. Thus for an engineering drawing distances in the

cardinal directions are represented faithfully, and The object can be reproduced accurately. However, diagonal and other off-axis distances are not displayed accurately.

Thus the isometric projection contains less information than a full set of orthographic projections. However, for many (if not most) people it is easier to visualize The 3-D shape of an object in the isometric display. Once again we are led to the situation where thought must be paid to how the information will be used.
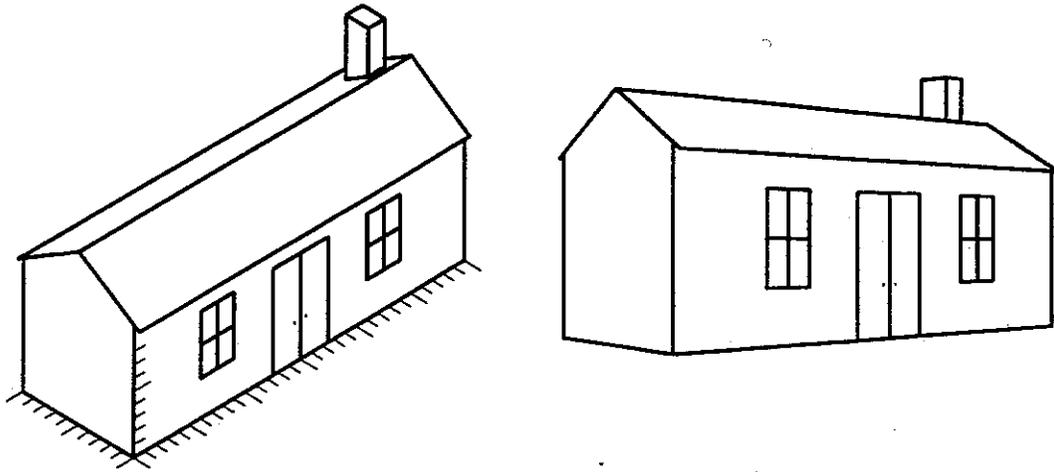


Figure 2-21 An isometric drawing (left), where correct distances can be measured to the same scale, in the three principal directions, and a perspective drawing (right), which looks natural but is not suited to measurement of dimensions.