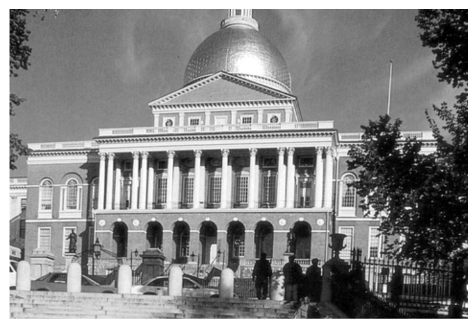
Where the Camera Was

KATHERINE McL. BYERS
JAMES M. HENLE
Smith College
Northampton, MA 01063
jhenle@math.smith.edu

How many times have you seen something like this?

Then

Now



Sources: Courtesy of The Bostonian Society/Old State House

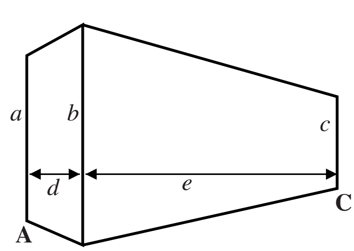
Simon Clay/Chrysalis Images

On the left is a picture of the Massachusetts Statehouse in Boston, taken about 1860. On the right is a picture taken in 1999. They appear in *Boston Then and Now* [**3**] and are meant to show us how the building and its setting have changed, but the effect is diminished because the camera was not in the same place for both photographs. How hard is it to determine the exact location of the photographer from information in a photograph?

The problem of understanding the relative positions of image and object is actively studied by computer scientists. In Kanatani [**2**], it is part of "computational projective geometry." The specific task of locating the camera from the photograph is called "camera calibration." In Kanatani's book the process is quite involved and technical. In a mathematical paper published later, Eggar [**1**] tackles the same problem. He proves that the task can be done, but the technique is similarly complex and the paper does not derive a practical method or formula.

In this paper, we present a method and a formula for locating the position of the photographer. Our basic result is the following:

PROPOSITION. *If a picture of a rectangular solid taken by a vertically-held pinhole camera has measurements (on the photograph) of a, b, c, d, and e,*



then the camera was positioned

$$\frac{dc}{d(b-c)+e(b-a)}\overline{\mathbf{BC}}$$

to the left of **B** *in the direction from* **C** *to* **B** *and*

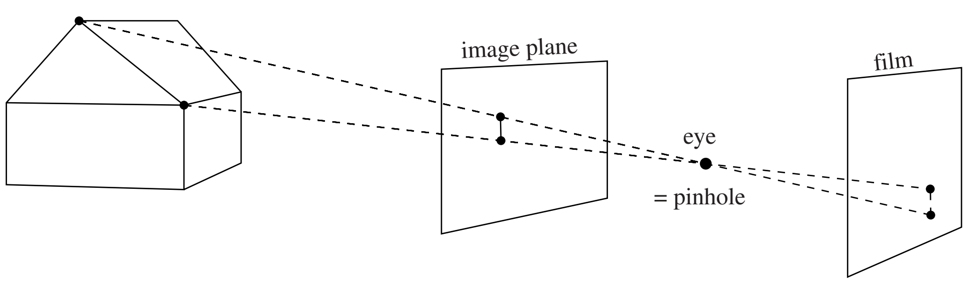
$$\frac{ae}{d(b-c)+e(b-a)}\overline{\mathbf{AB}}$$

in front of point **B**, *where* $\overline{\mathbf{BC}}$ *and* $\overline{\mathbf{AB}}$ *are on-site measurements.*

The proof is based on high-school plane geometry and the basic principles of projective geometry taught in a beginning drawing class.

Background

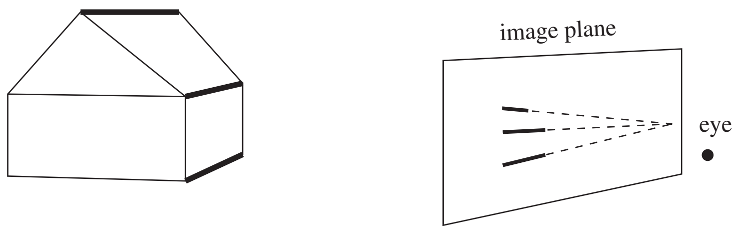
Our assumption is that the camera is a pinhole camera with the film in a vertical plane (plane perpendicular to the ground). Under these circumstances, the image on the film is the same as if we projected the three-dimensional world onto a plane, what we'll call the "image plane," using straight lines to the viewer's eye.



The only difference is that with a pinhole camera, the image appears on the film upside down.

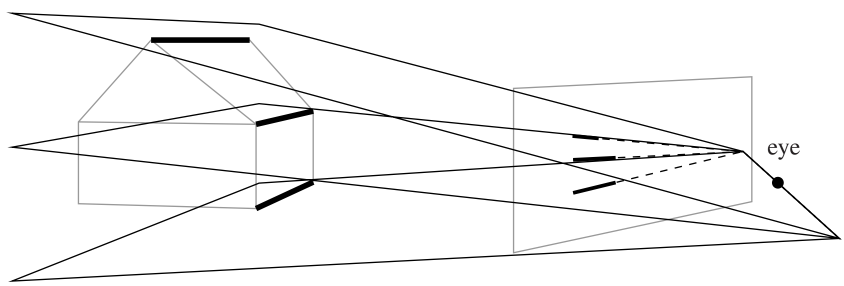
We'll need a few elementary facts about this projection:

(A) *The images of lines that are parallel to the ground and to one another, but not parallel to the image plane, meet at a single point in the image plane.*



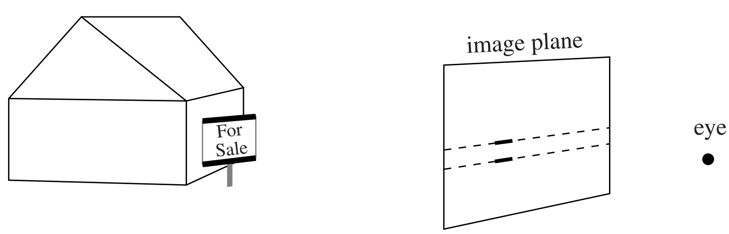
This point is called the **vanishing point** of the collection of parallel lines.

Imagine a collection of planes, each passing through the eye and one of the parallel lines. Then the planes intersect in a line that meets the image plane at the vanishing point.



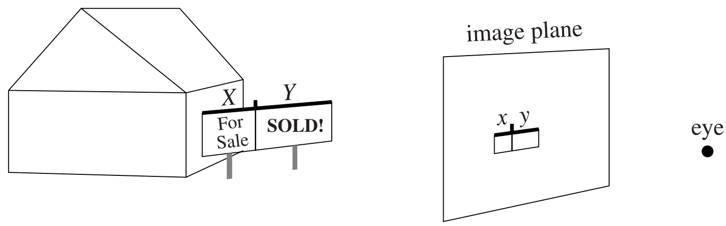
All such vanishing points lie on a single horizontal line called the **horizon line**.

(B) *Lines in the real world that are parallel to each other and also parallel to the image plane are parallel when projected onto the image plane.*



From this it follows that real horizontal lines are projected to horizontal lines.

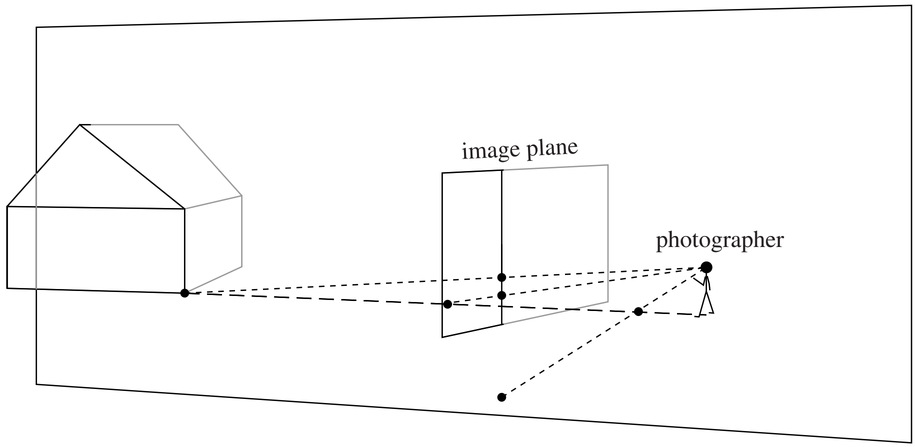
(C) *Also, ratios along lines parallel to the image plane are preserved when projected to the image plane. In the diagram below, this means that* $X/Y = x/y$.



Finally,

(D) *Lines on the ground connecting an object to the photographer appear as vertical lines on the image plane.*

Again, imagine a plane containing the eye of the photographer and the line to the photographer.



That plane is vertical and intersects the image plane in a vertical line.

A converse of **(D)** is also true: lines in the ground plane whose images are vertical connect to the photographer.

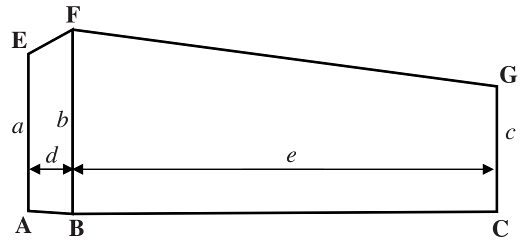
Our method

Given the tools above, we present a simple method for determining the location of the photographer.

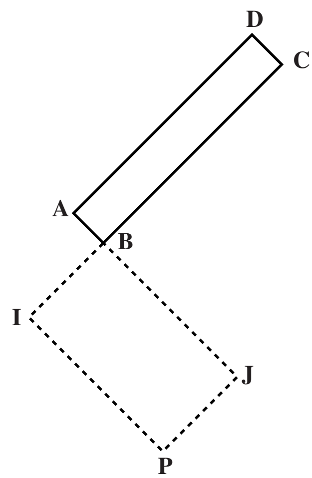
We start with a photograph of John M. Greene Hall at Smith College, taken around 1935 by Edgar Scott. Since the building is a complex solid, we pick a rectangular solid on it whose corners are easy to locate.

Source: Historic Northampton, Northampton, Massachusetts

We'll call this outline the *schematic picture*.



The schematic corresponds to the aerial view below, where **BC** is the front of the building and **P** is the location of the photographer.

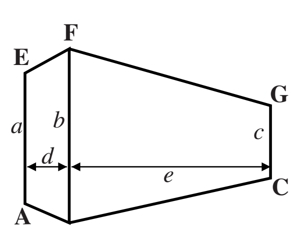


Our goal is to compute the distances $\overline{\mathbf{IB}}$ and $\overline{\mathbf{JB}}$. We'll compute $\overline{\mathbf{IB}}$—the computation of $\overline{\mathbf{JB}}$ can be done symmetrically. Our procedure is to express

$$\frac{\overline{\mathbf{IB}}}{\overline{\mathbf{BC}}}$$

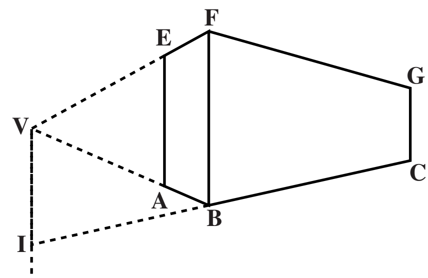
in terms of the five measurements a, b, c, d, and e in the image plane. Assuming we can measure $\overline{\mathbf{BC}}$ on site, we can then multiply this times the ratio to find $\overline{\mathbf{IB}}$.

To make the proof easier to view, we will show our work on a schematic with sharper angles:

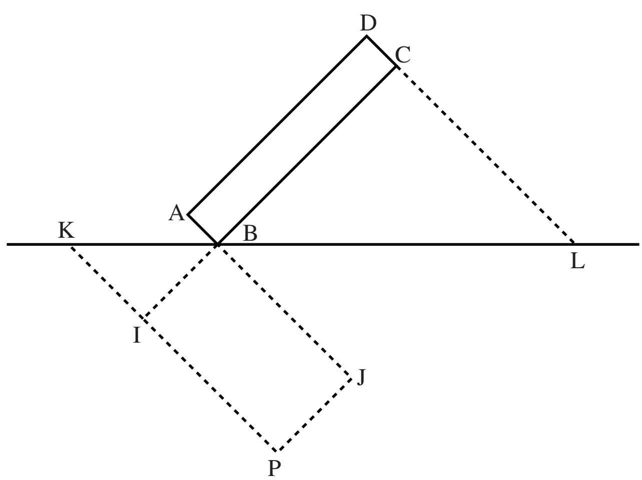


We begin by extending **EF** and **AB** in the schematic picture to determine the location of the left vanishing point, **V**.

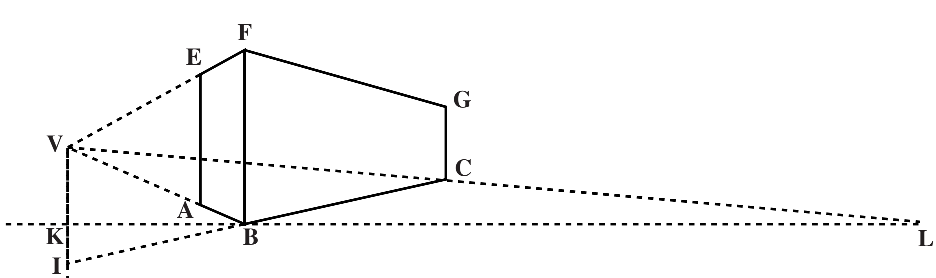
Next, notice that **PI** in the aerial view is parallel to **AB**, hence by Fact **(A)**, in the schematic picture it passes through **V**. Also, since it is a line to the photographer, by Fact **(D)** it is vertical in the schematic picture. Thus point **I** is the intersection of this vertical with the extension of **BC**.



Now we add a horizontal line through **B** parallel to the image plane and extend **PI** and **DC** to meet it. In the aerial view, it looks like:



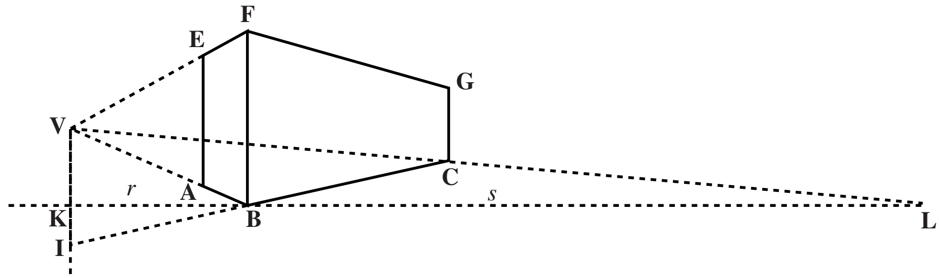
By Fact **(B)**, this line is also horizontal in the schematic. The aerial view line **CL** is parallel to **AB** and **PI**, so it too passes through **V**.



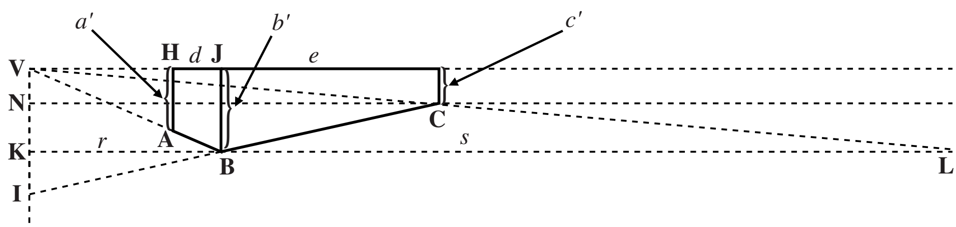
From $\triangle$**KIB** $\sim$ $\triangle$**LCB** in the aerial view we have

$$\frac{\overline{\mathbf{IB}}}{\overline{\mathbf{BC}}} = \frac{\overline{\mathbf{KB}}}{\overline{\mathbf{BL}}}.$$

From Fact **(C)**, this proportion is equal to the ratio of image plane distances r/s.



To find r/s, we add two more horizontal lines, **CN** and the horizon line **VH**, then focus on the lower half of the resulting figure.



From $\triangle$**VLK** $\sim$ $\triangle$**VCN** we have

$$\frac{r+s}{b'} = \frac{r+e}{c'}, \quad \text{from which we can derive:} \quad \frac{r}{s} = \frac{c'r}{b'r + b'e - c'r}.$$

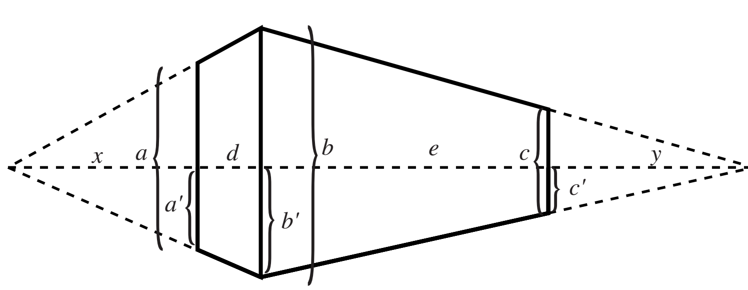
From $\triangle$**VJB** $\sim$ $\triangle$**VHA** we have

$$\frac{r}{b'} = \frac{r-d}{a'}, \quad \text{from which we can derive:} \quad r = \frac{b'd}{b'-a'}.$$

These together give us

$$\frac{r}{s} = \frac{c'\frac{b'd}{b'-a'}}{b'\frac{b'd}{b'-a'} + b'e - c'\frac{b'd}{b'-a'}} = \frac{c'd}{b'd + b'e - ea' - c'd}.$$

We promised to express this ratio in terms of a, b, c, d, and e. We can accomplish that by one more application of similar triangles: We have



$$\frac{a'}{b'} = \frac{x}{x+d} = \frac{a}{b}, \quad \text{and} \quad \frac{c'}{b'} = \frac{y}{y+e} = \frac{c}{b},$$

and so

$$\frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'}$$

giving us

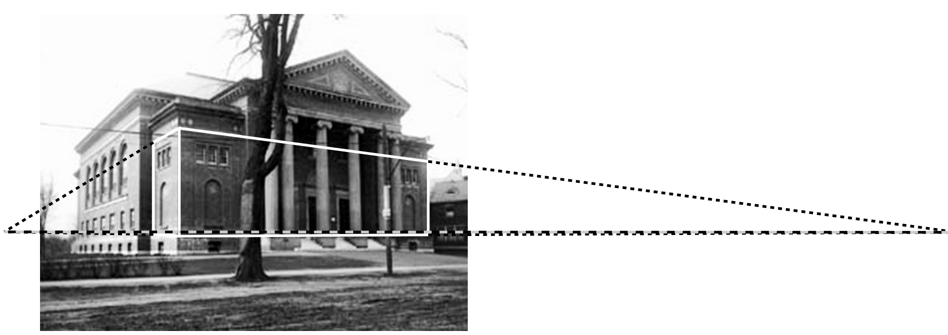
$$\frac{\overline{\mathbf{IB}}}{\overline{\mathbf{BC}}} = \frac{r}{s} = \frac{\frac{b}{b'}c'd}{\frac{b}{b'}b'd + \frac{b}{b'}b'e - e\frac{b}{b'}a' - \frac{b}{b'}c'd} = \frac{dc}{d(b-c) + e(b-a)}.$$

The corresponding formula for $\overline{\mathbf{BJ}}/\overline{\mathbf{AB}}$ can be found symmetrically:

$$\frac{\overline{\mathbf{BJ}}}{\overline{\mathbf{AB}}} = \frac{ae}{d(b-c) + e(b-a)}.$$

This completes the proof of the proposition. ∎

The last step in locating the position of the camera is finding its height. This is accomplished in a primitive way by noting where the horizon line cuts across the picture. The height of the camera is the height of this line as it appears against the building in the picture.



Source: Historic Northampton, Northampton, Massachusetts

Conclusion

The close agreement of the two pictures illustrates the proposition.

Then

Source: Historic Northampton, Northampton, Massachusetts

Now

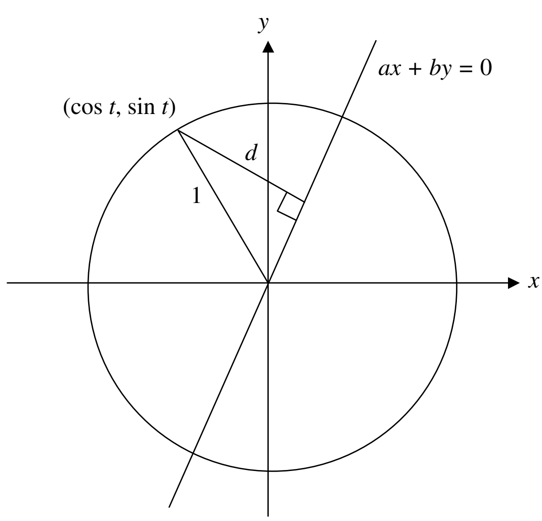
There are problems, though, in applying the proposition. It may be difficult to find an appropriate part of a building to analyze. It can be difficult to measure the building. It can be difficult to measure the photograph. Finally, locating the spot computed by the proposition, is not easy without equipment.

Considering these problems, the close agreement of the pictures of John M. Greene Hall might be considered good luck. We used a high-resolution scan on the archive photograph—b was measured at 470 pixels. Even so, if b were measured just one pixel less, the computed location of the photographer changes by almost two feet (because of the strategic location of b in the denominator of the formula).

REFERENCES

1. M. H. Eggar, Pinhole cameras, perspective, and projective geometry, *Amer. Math. Monthly* **105**:7 (1998), 618–630.
2. Kenichi Kanatani, *Geometric Computation for Machine Vision*, Clarendon Press, Oxford, 1993.
3. Elizabeth McNulty, *Boston Then and Now*, Thunder Bay Press, 1999.

Proof Without Words: Extrema of the Function $a \cos t + b \sin t$



$$d \leq 1 \Rightarrow |a\cos t + b\sin t|/\sqrt{a^2+b^2} \leq 1$$

$$-\sqrt{a^2+b^2} \leq a\cos t + b\sin t \leq \sqrt{a^2+b^2}$$

—M. Hassani, M. Bayat, and H. Teimoori
Institute for Advanced Studies in Basic Sciences,
P. O. Box 45195-159,
Gava Zang, Zanjan 45195, IRAN
Hassani@iasbs.ac.ir
Bayat@iasbs.ac.ir
Teimoori@iasbs.ac.ir