

Drawing on a Skating Rink

Douglas W. Dwyer

New York City, USA; douglas.dwyer@me.com

Abstract

We develop a methodology for placing line drawings within the three-dimensional space of an ice rink as captured by a photographer or an artist. We show that historical paintings and prints of ice skaters are consistent with an implication of rectilinear perspective. As a result, the same approach can be applied to historical artworks. The approach is contrasted with another algorithmic approach for placing many images in a frame and capturing a sense of depth. By making an assumption on the angle of view of the image, one can estimate the distance of different skaters from the camera. As an application, we draw multiple images in the space of the skating rink and animate them.

Introduction

Due to gravity and the size of the earth, an ice-skating rink can be viewed as a flat plane. In a level picture of an ice-skating rink with the horizon near the center, the plane of the rink should be approximately orthogonal to the plane of the picture. Can we use these hypotheses to discover how far away skaters are from the photographer? Can we draw additional skaters on the rink? Do historical paintings of ice-skating rinks adhere to the same rules of perspective as photographs?

There is an extensive literature on how to discover a three-dimensional space from a sequence of photographs taken over time. Such discoveries have many applications including the insertion of computer-generated imagery into a traditional movie (e.g., Match Moving). Specialized software and significant computer power is often deployed in making such discoveries. A subset of this literature studies the discovery of 3D space from a single photograph by using knowledge of the physical properties of the objects in the photograph (cf., [2]).

In the 1990s, the artist J. Seward Johnson was able to create sculptures that from a specific viewpoint replicate impressionist paintings [6]. His process included making a transparency of the painting, placing it on a tripod and then making the sculpture by viewing progress compared with the transparency from a specific viewpoint. Today, similar techniques are used to make 3D virtual representations of photographs with computer software.

This paper is a minimal application of such techniques. We discover a 3D space from a single photograph. After discovery of the space, one can place a “connect-the-dots” line drawings into this space without specialized software. We would like to make such techniques accessible to those that seek to utilize data analytic skills in the creation of art.

We also show that linear regression can be used to test the perspective of historical paintings. We show the test can be used to understand one historical painting as being the synthesis of two different viewpoints. We believe that this application is novel.

The Approach

The photograph in Figure 1 was taken by Sem Presser in about 1950 (© Sem Presser / MAI). The castle in the background is called Muider slot, which is in the Netherlands. In what follows, we will analyze imagery

using the open source statistical programming language called R. We will make use of the locator function to measure points on the photograph as well as the R package *magick* to read the photographs into R and plot text, lines and polygons on top of them.

In Figure 1, we plot the photo as well as the measured points. We have plotted the photo using *units* in which the y coordinate of the bottom of the photo is -1 and the top is 1 . The points locate the feet and the top of the heads of ten adults in the photograph. In what follows, we will refer to these units as *picture frame units* as they are relative to the top and bottom borders of the photograph.

There are photographs of Sem Presser (the photographer) holding a Rolleiflex 2.8 GX, which was a classic photojournalist camera of the period. It had a fixed 80mm lens and its film was 60mm x 60mm. Therefore, its angle of view is given by: $2 \arctan(60\text{mm}/(2 \times 80\text{mm}))$, which is about 41° (cf. page 51 of [7]). It appears that the photographer has composed the picture to get as much of the foreground in the picture as possible while at the same time not cutting off the top of the castle in the background—the photographer has rotated the camera downward. This is further indicated by the horizon line being above the center of the photograph and the lines of the castle converging towards a vanishing point below the horizon. We hypothesize that the photographer chose not to crop either the sky or the foreground—the vertical angle of view is 41° .

It is convenient to define an $R3$ coordinate system, where the camera sits at the origin, the x -axis is left to right, the y -axis is up and down, and the z -axis represents the distance from the camera where a positive number indicates how far in front of the camera the point resides. We will define the picture plane as the set of points in which $z = 1$. This system allows one to project a set of points onto the picture plane by simply dividing their x and y coordinates by their z -coordinate, because the intersection of a line defined by the origin and a point $\{x, y, z\}$ with the plane defined by $z = 1$ is $\{x/z, y/z, 1\}$. To distinguish between points in 3D space and points on the picture plane we will refer to the latter as $\{u, v\}$, which are the x and y coordinates in the picture plane, respectively.

Figure 2 presents a side view this coordinate system. From the side, the vertical axis is the y -axis and the horizontal axis is the z -axis. The origin as well as the points $\{(0, -1), (1, -1), (2, -1 + 2\delta), (3, 3\delta)\}$ are indicated. The ice rink is represented as a white line that intersects with the vertical axes at -1 and has a slope of δ . The picture frame is indicated with a gold vertical line in the $z = 1$ plane. The top and the bottom of this line are determined by the angle of view assumption that we use. The picture dots on the picture plane and the corresponding black dots on and above the ice rink represent the bottoms of the feet and the tops of the heads of the people in the photograph that we measured. The points are connected with the origin by light grey lines which indicate that the camera is located at origin. A white line that is labeled ‘horizon’ goes through the origin and is parallel to the ice rink. This line would connect a person that is arbitrarily far away with the origin. The intersection of this line with the picture frame indicates the location of the horizon line in the photograph. We have also added the side views of the drawings that will be placed into the photograph in the Artistic Application that follows. We can think about a rotation downward of the camera as the rink having a positive incline relative to the picture plane. Therefore, the plane of the ice rink can be defined as the set of all points in which the y coordinate is equal to $\delta z - 1$, which assumes that the rink is 1 unit below the camera when z is 0. The construct δ will play



Figure 1: Ten people in the picture plane

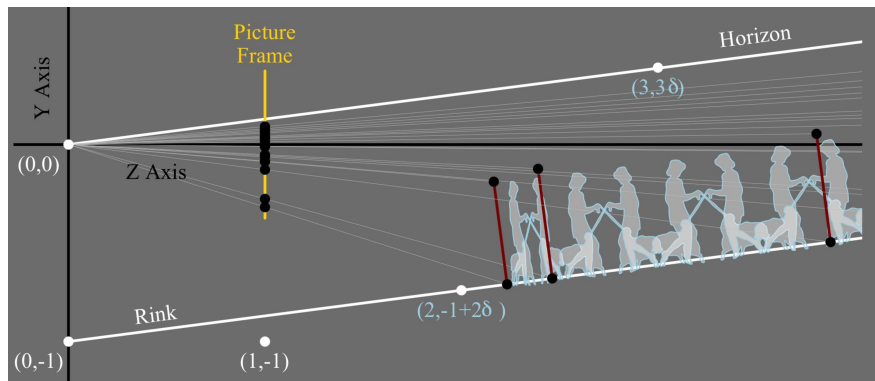


Figure 2: *Side View*

a key role in what follows: it is the slope of the ice skating rink relative to the picture plane, and it is also the location of the horizon line on the picture plane at $z = 1$.

Using this system, we can represent the location of the i th person's feet in R^3 using a function that returns each coordinate given the x and z coordinates of the i th person's feet:

$$xyz(i; F) = \{x_i, \delta z_i - 1, z_i\}$$

In this *coordinate system*, we have normalized the height of the camera to 1, and we refer to the units as normalized units to contrast them with picture frame units. The parameter F indicates that the function returns the location of the bottom of the person's feet. The location of the i th person's feet in the picture plane can be represented as a similar function:

$$uv(i; F) = \{x_i/z_i, \delta - 1/z_i\}$$

Further if h_i is the height of the i th person, the position of the top of their head in R^3 and on the picture plane is given by:

$$xyz(i; H) = \{x_i, \delta z_i - 1 + \cos(\phi)h_i, z_i - h_i \sin(\phi)\}, \text{ and}$$

$$uv(i; H) = \{(x_i)/(z_i - h_i \sin(\phi)), \delta + (-1 + \cos(\phi)h_i)/(z_i - h_i \sin(\phi))\}.$$

If $h_i \sin(\phi)$ is small relative to z_i we can write:

$$uv(i; H) \approx \{x_i/z_i, \delta - (1 - \cos(\phi)h_i)/z_i\}.$$

where the parameter H indicates that the function is returning the location of the top of the person's head. Note that we use $\sin(\phi)h_i$ and $\cos(\phi)h_i$ to account for each person being pitched towards the picture plane due to the downward rotation of the camera. The angle of inclination, ϕ , is given by $\arctan(\delta)$. The approximation error is small when people are far away from the picture plane and the downward rotation of the camera is small.

From this relationship, we see that a person that is very far away will be a *dot* on the horizon where the y coordinate of the horizon line is δ in the picture plane. Given the vertical coordinate of the location of a person's feet on the picture plane, we can solve for the z_i coordinate of their position on the rink: $z_i = -1/(v_i - \delta)$. In the picture plane, the vertical distance between the feet and the head of the i th person, vh_i , is given by taking the difference between the v coordinates in the picture plane:

$$vh_i = \cos(\phi)h_i/z_i = -\cos(\phi)h_i(v_i - \delta)$$

In the second expression we substituted in v_i for z_i , which implies a linear relationship between the projective height of a person and the distance below the horizon of person's feet in the picture plane. We can estimate this linear relationship in picture frame units (what we observed) using linear regression. The result of the regression can be used to place the observed points into the picture plane.

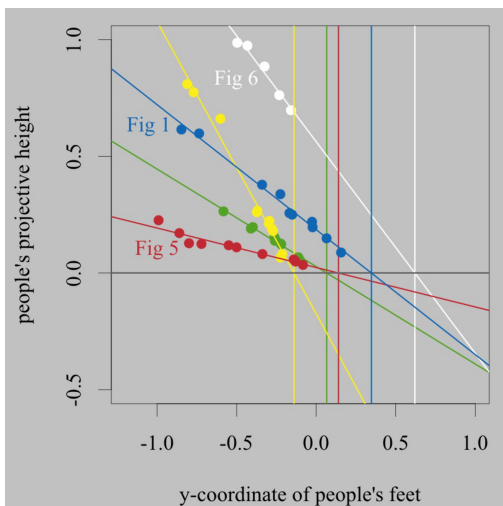


Figure 3: Scatter plots for five images

The red dots in Figure 3 presents the scatter plot implied by the above equation for Figure 1. The horizontal axis is the v coordinate of each person's feet in the projective plane, and the vertical axis is the projective height of each person. We draw the *best fit* line through the points, and we can solve for the v coordinate of a person who would have 0 projective height in the picture plane. It is 0.347, which is the negative of the intercept divided by the slope of the best fit line and yields an estimate of where the horizon (δ) is in picture frame units. We can use this estimate to transform the observations in the picture into the normalized units that we use in $R3$. This transformation allows one to solve for the $R3$ location of our 20 points.

In Figure 2, the top and bottom of the gold vertical line represents the picture frame. The vertical coordinates of the top and bottom of the picture frame are $\pm \tan(41^\circ/2)$ in normalized units. The location of the horizon is δ which we measure to be 0.347 in picture frame units. We can transform this into normalized units by multiplying by $\tan(41^\circ/2)$. This implies a δ of 0.130 and angle of incline of about 7.4° . Therefore, we can transform the 20 points that we measure in the picture frame into the normalized $R3$ units by multiplying them by $\tan(41^\circ/2)$.

This transformation allows us to solve for the location of each person's feet and each person's height in the picture plane in normalized units as well as the location of each person's feet on the ice-skating rink and the height of each person (in $R3$). For the average height of each person, we get 0.31. This implies that the height of the camera is about 3 times the height of the average person in the picture. The photographer must have been standing on something that allowed him to look down on the skating rink. Further, if we assume that the average height of the people in the picture is 5'8", we get that the closest person is 24 feet away and that the furthest person is 150 feet away when measured along the z -axis.

Artistic Application

Once we have the coordinate system of the ice rink, we can make line drawings of people on the photograph so that their sizes are consistent with the corresponding people on rink in the photograph.

In 2023, we developed a methodology to create a "connect the dots" 2-D representation of a silhouette [4]. We can place this two-dimensional representation into a 3D space by transforming it such that the height of the image is comparable to the average height of the figures in the photograph, moving it to the desired location and rotating as required.

For example, we draw a woman walking her dog and repeat it as if she is walking a *figure-eight* around the ice rink as follows. We first place a rectangle on the rink that encompasses most of the central people on the rink, we then transform a parametric representation of a figure-eight, $\{\sin(2\theta), \cos(\theta)\}$, to fit within this



Figure 4: *Drawing placed multiple times on photo*

region.

We next choose one point on the figure-eight and place a drawing of a woman walking a dog such that the bottom of the image aligns with the chosen point on the figure-eight. We scale the drawing such that height of the woman is comparable to the average height of the figures in the photograph. We rotate the drawing using *yaw* about its vertical axis so that the drawing is in a vertical plane that coincides with the line that is on the plane of the rink and tangent to the figure-eight at the chosen point. We then project the figure into the picture frame by dividing the x and y coordinates by the z coordinate. We plot the drawing with a polygon using a transparent color and a boundary on top of the photograph. We repeat using a set of points that are evenly spaced around the figure-eight enough times so the images almost touch (Figure 4).

We found that the illusion could be enhanced by pushing the image slightly below the surface of the rink and masking out certain portions of the drawings to allow some of the figures in the photograph to be in front of the drawn figures. Otherwise, the drawings appeared to float over the people in the photograph in the distance. The repetition of the dog walker suggests the women and the dog walking around the rink along the figure-eight pattern. The image can be animated as desired.

Do Historical Paintings Obey These Rules?

We have a prediction that the projective heights of people on a plane that is inclined towards the picture plane will decline linearly as they approach the horizon line. Many pre-photography images of ice-skating rinks obey the rules of rectilinear perspective. Figure 3 presents five scatter plots derived from four additional images (one is a photograph, one is a print, and two are paintings) using a different color for each image.

Three of the images (Figures 1, 5 and 6) are in this paper, and the other two are available in a supplement to this paper. In all cases, the “fit” of the regression is reasonable and for all but one, the estimated horizon line is consistent where we would have “drawn it” on top image had we been asked.

Figure 5 applies this approach to the painting “Winterlandschap met schaatsers,” by Hendrick Avercamp from about 1608. The numbers indicate the location of the feet and the heads of each person that was used in the analysis. The brown horizontal line drawn over the image is the estimated horizon line, which is consistent with where we would have drawn it. Using 41 degrees for the angle of view, we find the closest person is 78 feet away and the furthest person is 396 feet away (along the z axis). The viewpoint is estimated to be 35 feet above the plane of the ice rink. A box has been placed on the plane of the skating rink and an image of a couple has been placed in the corners of the box as well as the center. The height of the couple aligns with the corresponding heights of figures in the actual painting (the figures whose feet have approximately the same v -coordinate in the picture).



Figure 5: *Figures on Avercamp’s Winter Landscape.*

In fact, the regression line estimates a horizon line that is just above the top of the image, which is drawn with a brown horizontal line (see the white dots and line in Figure 3). Once again, a box has been drawn on the plane of the rink and a figure has been drawn five different times so that the figure’s height is consistent with the corresponding figures in the image.

The artist, in contrast to the estimated horizon line, has indicated a horizon line that approximately divides the top third of the picture from the bottom two thirds. One could interpret this as the artist has chosen to combine two points of view into one image. The ice-skating rink itself is composed such that the horizon line divides the picture. The figures are composed such that they are viewed from approximately eye-level. The painter David Hockney has written about how painters would combine different points of view into the same picture in European painting during this time period and more generally (cf. page 152 [5]).

Other Approaches

In Bridges 2023, Demaine, Demaine and Bass presented an algorithm for arranging a collection of images that achieves a number of objectives [1]. All images fit within the frame, and the smallest image is still visible. A sense of depth and height is conveyed. The random placement of images has a natural feel – not too bunched together. Their algorithm works by making the images at the top of the frame smaller relative to the images at the bottom of the frame and plotting the images at the bottom of the frame on top of the images at the top of the frame after a random placement of each image. They applied their algorithm to produce a sculpture using origami with *curved creases* that contained exactly 41,732 images of people, which was the 2021 population of Fitchburg, Massachusetts.

Their paper explicitly links the scale of each drawing to the y -coordinate of the drawing, where y can be viewed as the distance below the horizon in the picture plane. They use two equations: $\zeta_1 = \eta^{1.5}$ and $\zeta_2 = \eta^3$, where ζ_1 determines the image's scale, and ζ_2 determines the image's y coordinate and η is a "common number uniformly distributed between 0 and 1" (See section 'Appearing Uniform' of [1] and note that we have replaced z with ζ and \bar{z} with the η).

In the context of this paper, ζ_1 can be viewed as the projective height of the image (or h/z) and ζ_2 can be viewed as negative of the y -coordinate in the picture plane (or $-y/z$). If we normalize the scale of each image to one by setting h to 1 in 3D space, we can solve for the z and y -coordinates that would be consistent their chosen parameters: $z = 1/\eta^{1.5}$ and $y = -\eta^{1.5}$, which imply: $y = -1/z$. Therefore, one could replicate the algorithm of [1] in 3D space by placing each equaling sized drawing on top of a "hyperbolic slope" and positioning the camera accordingly.



Figure 6: *Figures on Avercamp's Skaters on Ice*

In making a large group photo, a photographer (or portrait painter) can position each person to achieve objectives similar to those of [1]. Photographers may choose to use a staircase or a sloped hill to compose a large group photo rather than having everybody stand on a flat plane, so that the subjects that are further away are elevated. By doing so, the photographer slows the rate at which people become smaller as they move up the picture plane (relative to everyone standing on a flat surface), and as a result it is easier to get everyone in the picture. The parameterization that [1] arrived at to achieve a *pleasing distribution* of many images in one frame is similar to what photographers do when they use a sloped hill or staircase to make a large group photo.

Our System of Moving from 3D to 2D in Context

Projecting three-dimensional space into a two-dimensional picture plane goes back hundreds of years. Our system of defining the set of points in which $z = 1$ as the picture plane thereby allowing a point in 3D space to be projected into the picture plane by dividing through by z was inspired by Carlbom and Paciorek's Figure 3-21 [3]. Our system is non-conventional in that it does not obey the *right-hand rule* for axes orientation. To be consistent with this rule, the distance *in front* of the camera would be a negative number and the picture plane would be the set of points defined by $z = -1$.

A related paper determines measurements in a single picture from a vanishing line and an additional vanishing point [2]. Their methodology includes solving for multiple vanishing points using multiple sets of parallel lines in the picture. The parallel lines are often edges of buildings. Our methodology, in contrast, assumes that a set of people are standing on a plane and determines the vanishing line (the horizon) through regression analysis. Our methodology works in the absence of sets of lines known to be parallel, although we do use an angle of view assumption.

The Role of Angle of View

One cannot determine the angle of view directly from the points in a picture plane. If one has a set of points in 3D projected onto a picture plane and transforms the z axis by a constant factor, λ , then one has just effectively rescaled the projection by the inverse of λ , and one can get the same image by rescaling the picture

plane. If one could see a side view of the new set of points, one would see that the image is distorted but the distortion would be invisible to the viewpoint of the camera.

For purposes of measuring the incline of the skating rink, the distance to the closest and furthest person and the distances between them, the result is sensitive to the angle of view assumption. As the angle of view assumption becomes smaller, the measured incline of the rink decreases and the figures become further away, but the ratio of the distance between the furthest subject and the closest subject is invariant to the angle of view assumption.

By assuming that the top (bottom) of the picture frame is the tangent of one-half (minus one-half) the angle of view, we are implicitly assuming that the picture frame is centered on the picture plane. Most cameras and lens are designed to work this way. Nevertheless, this assumption may be incorrect in images that have been cropped. Also, some lenses are known to produce a distorted image. The approach could be extended to handle such issues as required.

Conclusion

In this paper, we have shown how to identify the coordinate system of a historical photograph or painting that feature skaters on an ice rink by using linear regression. We have shown how images can be placed into the space such that they look like they are standing within the three-dimensional space of the photograph. With a angle of view assumption, we have shown how to estimate distances to the different people in the image as well as the height of the camera above the rink. We have shown that the approach can be used to assess whether or not a historical painting is consistent with the rules of perspective. We have identified one painting that appears to synthesize two different points of view using an application of this approach.

The same approach could be applied to other settings in which people stand on a surface that is known to be approximately flat, such as a playing field, a town square, or people standing at the edge of a body of water. A gallery supplement provides two more examples with links to animations. One future application of the approach could be to animate one person moving between two different images.

We thank the anonymous reviewers for the helpful and detailed comments.

References

- [1] M. Bass, E. Demaine and M. Demaine, “Algorithmic Layout of Characters in Perspective,” *Bridges Conference Proceedings*, Halifax, Nova Scotia, Canada, July 27-31, 2022, pp. 5–14. <http://archive.bridgesmathart.org/2023/bridges2023-5.html>.
- [2] A. Criminisi, I. Reid, and A. Zisserman, “Single View Metrology”, *International Journal of Computer Vision* November 2000, 40, pp. 123–148.
M. Bass, E. Demaine and M. Demaine, “Single View Metrology”, *Bridges Conference Proceedings*, Halifax, Nova Scotia, Canada, July 27-31, 2022, pp. 5–14. <http://archive.bridgesmathart.org/2023/bridges2023-5.html>.
- [3] I. Carlbom and J. Paciorek, “Planar Geometric Projections and Viewing Transformations”, *Computing Surveys*, Vol 10. No. 4, December 1978
- [4] D. Dwyer. “Drawing with Statistics”, *Bridges Conference Proceedings*, Halifax, Nova Scotia, Canada, July 27-31, 2022, pp. 157–165. <http://archive.bridgesmathart.org/2023/bridges2023-157.html>.
- [5] D. Hockney. *Secret Knowledge: Rediscovering the Lost Techniques of the Old Masters*, Viking Studio, 2001.
- [6] S. Johnson. *Beyond The Frame: Impressionism Revisited*, Bulfinch Press, 2003.
- [7] Y. Ma, S. Soatto, J. Kosecka, and S.S. Sastry. *An Invitation to 3-D Vision from Images to Geometric Models*, Springer Science and Business Media, 2004.