

1 Signed Distance Functions

1.1 Theory

maybe better if definition are highlighted

As Hart says in the introduction of “*Sphere tracing: a geometric method for antialiased ray tracing of implicit surfaces*” [3], an implicit surface is defined by a function that, given a point in space, indicates whether the point is inside, on or outside the surface. The implicit surface can be defined either with an algebraic distance or with a geometric distance. As an example, the unit sphere can be defined with the algebraic implicit equation

$$x^2 + y^2 + z^2 - 1 = 0$$

or geometrically with

$$||x|| - 1 = 0$$

where $x \in \mathbb{R}^3$ and $||\cdot||$ is the Euclidean magnitude. In this document we will not analyze the differences between the two representations because we will use simple surfaces, for which the equations can be translated easily from one form to the other. Note that the two equations have the same value, i.e. zero, when the provided point belong to the surface, and both became negative when the point is inside the surface.

Formalizing, we can say that a *distance surface* is implicitly defined by a function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ that characterizes $A \subset \mathbb{R}^3$, set of points that are on or inside the implicit surface:

$$A = \{x : f(x) \leq 0\}$$

The surface can be also defined with $f^{-1}(0)$, which gives exactly the points on the surface. Even if the continuity of f and its negativity on the interior of the surface are not strictly necessary properties, they come useful in practice so, when possible, they are preferable. We can define the surface from the outside using a *point-to-set* distance:

$$d(x, A) = \min_{y \in A} ||x - y||$$

Thus $d(x, A)$, given a point x , returns the shortest distance to the surface defined by A . We say that $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ is a *signed distance bound* of its implicit surface $f^{-1}(0)$ if and only if

$$|f(x)| \leq d(x, f^{-1}(0))$$

Here we are saying that f is always at least as cautious as the true distance function $d(x, f^{-1}(0))$. This mean that we can move x by $f(x)$ in every direction and in the worst case (or best) we'll just hit the surface. Using f we'll never get to the interior of the surface. When the equality holds, f is a *signed distance function* (SDF) and returns the exact distance to the surface, so moving x by

this distance (in the right direction) means putting it exactly on the implicit surface.

In the introductory article “*Rendering Implicit Surfaces and Distance Fields: Sphere Tracing*”[4] by scratchapixel.com, these concepts are emphasized calling them respectively: *distance underestimate (implicit) functions* (DUFs) and *distance implicit functions* (DIFs). Remembering that $DUF(x) \leq DIF(x)$ for every DUFs and DIFs that respect the definitions above.

For simple shapes we can derive SDFs by hand, but when it’s not feasible we can use the Lipschitz constant λ , as Hart suggested [3]. We say that a function $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ is *Lipschitz over a domain D* if and only if there is a positive, finite, constant λ such that

$$|f(x) - f(y)| \leq \lambda \|x - y\|$$

Such λ is called the Lipschitz constant. Of course there is no upper limit to λ , but there is a lower bound; the function $Lip(f)$ return such minimum value. Reworking the formula we can observe that

$$\begin{aligned} \lambda &\geq \frac{|f(x) - f(y)|}{\|x - y\|} \\ \lambda &\geq \frac{f(x) - f(y)}{\|x - y\|} \\ \lambda &\geq \lim_{\|x - y\| \rightarrow 0} \frac{f(x) - f(y)}{x - y} = f'(x) \end{aligned}$$

the last step, given the continuity of f , permits us to estimate a safe lower bound for the Lipschitz constant. To be precise what we have to do is to compute the derivative maximum value, therefore we need to find the solutions to the function second order derivative $f''(x)$.

The following theorem (from [3]) explains why having such a λ (or a suitable approximation) is so important: because it gives us a general method for computing a DUF for any Lipschitz function.

Theorem 1.1. *Let f be Lipschitz with Lipschitz constant λ . Then the function f/λ is a SDF of its implicit surface.*

Proof. Given a point x and a point $y \in f^{-1}(0)$ such that

$$\|x - y\| = d(x, f^{-1}(0)) \tag{1}$$

Because f is Lipschitz, holds

$$|f(x) - f(y)| \leq \lambda \|x - y\|$$

and because y is on the surface $f(y) = 0$, so

$$\begin{aligned} |f(x)| &\leq \lambda \|x - y\| \\ |f(x)| &\leq \lambda d(x, f^{-1}(0)) \quad \text{by Equation 1} \end{aligned}$$

Hence $\frac{|f(x)|}{\lambda} \leq d(x, f^{-1}(0))$ is a signed distance bound (DUF) for any Lipschitz function f . \square

One secondary thing to note is that an SDF can return any negative number when the given point is inside the surface. A common choice is to give a constant negative number i.e. -1 .

1.2 Examples

Hart’s method is a general approach but in practice we can create by hand several DIFs. Different SDFs can be found in both [3, 4], but one of the most comprehensive lists is at [2] by Inigo Quilez, creator of shadertoy.com and “SDF artist”.

Below are reported three simple images generated with a modified version of the scratchapixel’s code[4]. Pixels’ color depends on the absolute distance¹ from

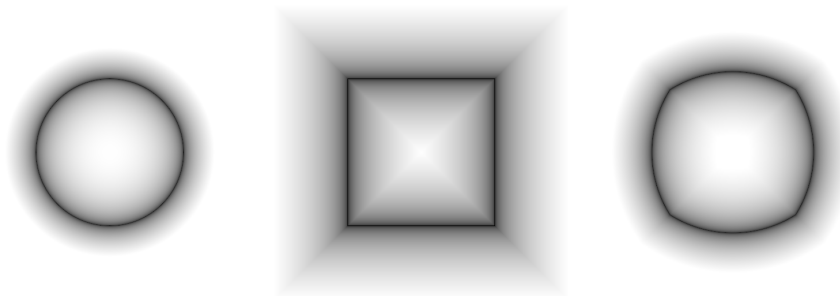


Figure 1: A circle

Figure 2: A square

Figure 3: A squircle

the surface. Note that this images can be seen as 2D section of 3D objects done with a yz -aligned plane² passing through the origin. The distance equation for the circle is $x^2 + y^2 - 1$, meanwhile for the square it is $\max(|x|, |y|) - 1$. To get a better understanding of the square’s SDF we suggest this Quilez’s video [1] (and the others on his channel). The squircle is the mixing of the square and circle SDFs and its implicit equation is $k * (\max(|x|, |y|) - 1) + (1 - k) * (x^2 + y^2 - 1)$, where $k \in [0, 1]$. This is only one example of the operation we can perform on implicit surfaces.

1.3 Constructive Solid Geometry

One of the major benefits we get from using SDFs is the easiness with which we can create complex shapes from few primitives, technique known as Constructive Solid Geometry (CSG). CSG with Boolean operations is largely used in CAD software. With DIFs we can simulate:

- union with the minimum, because we “stop” at the first surface that as been found;

¹gamma corrected and clamped in $[0, 1]$

²assuming that x grows to the right, y up, and z “towards” the image

- intersection with the maximum, because we “ignore” the first surface if there’s another after;
- subtraction with the maximum between the first surface and the opposite of the second

$$\max(f_1, -f_2)$$

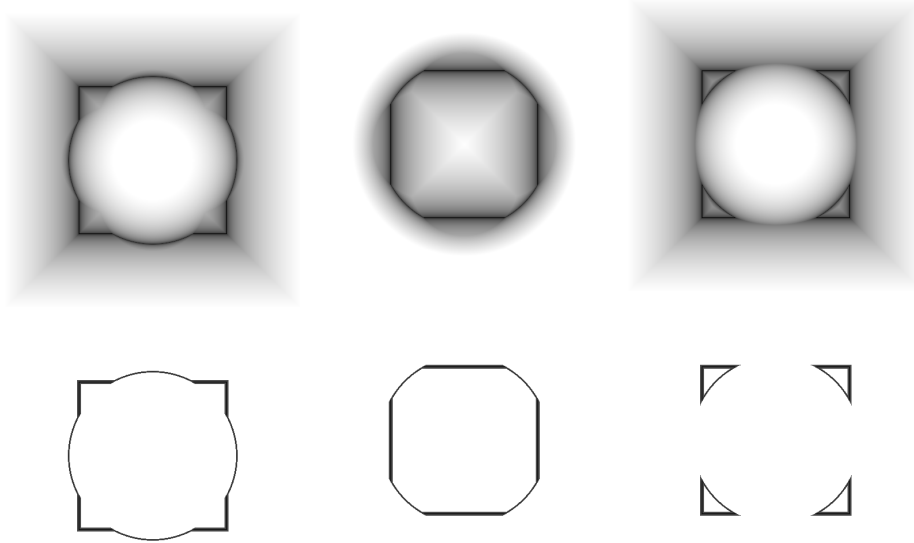


Figure 4: Square-Circle union and its contour

Figure 5: Square-Circle intersection and its contour

Figure 6: Square-Circle subtraction and its contour

Note that the signed distance bound functions are closed by CSG operations so we do not have to handle new cases.

Mixing and Blending and "all functions work" (modulo example)

2 Ray Tracing

In the introduction of the book “Ray Tracing from the Ground Up” [5] the author loosely defines two classes of renderers: projective algorithms and image-space algorithms. The former nowadays is the most popular because it’s based on the graphics pipeline model and can easily take advantage of modern graphics cards. The basic idea is to project the objects on the scene into a 2D plane and then compute the appearance of the objects that can be seen. The image-space algorithms start from the pixel and try to figure out where the light came from for that pixel. Following the light’s path we can compute hard shadows, soft shadows, reflection, refraction, depth of field, motion blur, caustics, ambient occlusion, indirect lighting and other effects. But following every light ray’s bounce it’s extremely costly, especially if there are a lot of light sources in the scene, and that’s why this approach historically wasn’t the one to be used for interactive application. Instead, given the high quality and fidelity that this method can achieve, it’s been widely used to create movies, clips and images, because in these application, after an initial low-res prototyping, the computer can elaborate the final product even for days.

say that the rt is being used also in real time application, blended with standard pipeline

2.1 Rays and Cameras

As we said before, the idea behind a ray tracer is to generate a ray for each pixel and check whether the ray hits a surface. If so it has also to find out if a light source lighten the point or if it is in shadow.

The easiest way to generate such rays is the following: imagine to put a 2D rectangle into a 3D scene; draw a square grid on the rectangle and for each square project rays perpendicular to the rectangle’s faces; if a ray hits an object color the corresponding square with the surface color; if a ray hits no objects then color the pixel with black (or a default color). Because the rays are parallel to one another this method gives us a really simple orthogonal camera.

Orthogonal cameras are useful but in general prospective cameras are preferred because they resemble more the human eye. To model a perspective camera all we have to do is to put a point behind the rectangle with the grid and generate rays that start from the point and pass through the grid’s squares. The common choice is to make the rays pass through the squares’ center. The model just described is called a pinhole camera, shown in Figure 7, because mimics the real life cameras. One trick here is that we put the plane before the point so we don’t have to flip the image before showing it to the user.

A ray is defined as a point and a direction, as shown in Figure 8, in formulas:

$$r(t) = O + tR$$

where $O, R \in \mathbb{R}^3$ and R is a unit vector. At each time t we can compute a point position on the ray, in this way the ray tracing algorithm can check for each t

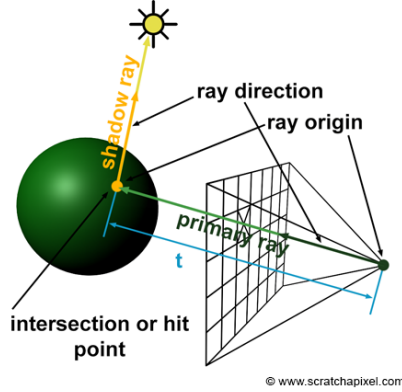


Figure 7: A pinhole camera inside a scene with a sphere and a light source

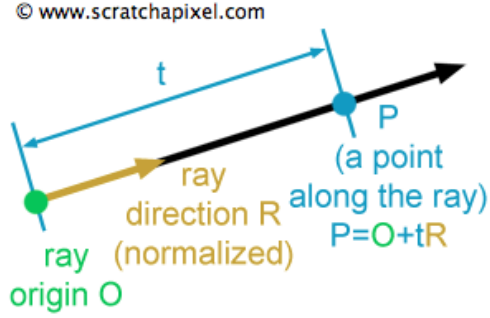


Figure 8: Ray model

whether the ray intersected a surface or not.

When generating rays we have to take in consideration several factors: camera position and orientation are the more obvious, but also the field of view and the image aspect ratio. First of all we need to derive the rays' direction with the pixel's centers on the image plane (the rectangle with a grid) and the camera origin (also called *eye*). But these two values refers to different coordinate systems because the pixel position refers to the top left corner of the image plane and the camera origin is defined in world coordinates. Here with world coordinates we mean the coordinate system with which the object (and the camera) are placed on the scene. One way to solve this problem is to produce the rays as if the camera was placed in the origin looking towards $-z$ and then translate and rotate the rays according to the true camera position and orientation.

In order to do so we need to remove the dependency from the actual output image size. Given two coordinates in raster space we normalize them by dividing respectively by the image width and height. Now the coordinate are in Normalized Device Coordinates (NDC) space. Actually the precise operation is the following, in which we also move the coordinate in pixels center:

$$PixelNDC_x = \frac{(Pixel_x + 0.5)}{ImageWidth}$$

$$PixelNDC_y = \frac{(Pixel_y + 0.5)}{ImageHeight}$$

Now we have to map this coordinates from $[0, 1]$ into $[-1, 1]$ because the eye of the camera is at the origin, so the center of the image plane (also called screen) is on the z -axis. In practice we are aligning the sign of the four quadrants of

the xy -plane in world coordinates with the image plane.

$$\begin{aligned} PixelScreen_x &= (2 * PixelNDC_x - 1) * AspectRatio \\ PixelScreen_y &= 1 - 2 * PixelNDC_y \end{aligned}$$

Remember that we want a right handed system so we map: the top left corner, $(0, 0)$ in raster space, to $(-1, 1)$ in screen space; the bottom right corner, $(ImageWidth, ImageHeight)$ in raster space, to $(1, -1)$ in screen space; and the last two accordingly.

The *AspectRatio* used in the formula equals $ImageWidth/ImageHeight$ and we must consider it when the output image is not squared. In fact we need to produce more pixel in one direction otherwise the generated image will look squished.

One of the last thing to consider is the field of view (fov), the view angle of the camera. Assuming that the distance between the camera eye and the image plane equals to 1, we want to know by how much enlarge or shrink the image plane horizontally and vertically. Knowing the fov angle α we just need to multiply the screen coordinates by $\tan(\alpha/2)$.

The last thing to take in consideration is that our camera may not be always positioned in the world origin and look toward $-z$. To do so there are several options the one explained here is also

explain like if we are moving a point with a normal direction

2.2 Sphere Tracing

Introduced by Hart[3] in 1996, the sphere tracing algorithm gives a better alternative to other ray tracers like ray marching. In ray marching each ray is analyzed by constant steps of t . We can detect a ray-surface contact using the sign of the distance function: when the ray enters the surface the distance became negative. Let's say that for a ray r , $d(r(t_k), S) > 0$ and $d(r(t_{k+1}), S) < 0$, for some instant t_k and the following instant t_{k+1} and an implicit surface S . Then using bisection on the segment between the two points $r(t_k)$ and $r(t_{k+1})$ we look for the point P on the segment for which $d(P, S) = 0$. To actually find this point P on the surface S can be costly, so usually it's found a point that is as near to the surface as a given threshold.

We can speed up the algorithm by increasing the step size, i.e. $t_k - t_{k+1}$, but doing so leads two problems: performing bisection is slower because the segments are longer; if an object is small the algorithm may evaluate the ray just before and after the object, thus not seeing it. While the first problem can be managed, the second can cause completely wrong images to be produced. What we would like is an algorithm with an adaptable step size, to go fast when there are no surfaces near and to slow down when there are objects in the proximity: sphere tracing does that.

Remember that before we said that $d(x, S)$ for some $x \in \mathbb{R}^3$ means we can move x in every direction by $d(x, S)$ being sure at worst to just hit the surface. In other words we are drawing a safe sphere around the point, in which there

is no surface. From the surface perspective this sphere is called an *unbounding sphere*[3], and the union of all the unbounding spheres describes the surface because no sphere penetrate the object. The concept is approximately the opposite of *bounding volume*, a volume that surrounds the object, because here we define an area of space not containing the object. In sphere tracing the ray-surface intersection is determined by a sequence of unbounding spheres, as shown in Figure 10.

```

t = 0
while (t < D) {
  d = f(r(t))
  if (d < ε) {
    // intersection
    reutrn t
  }
  t += d
}
// no intersection
return ∅

```

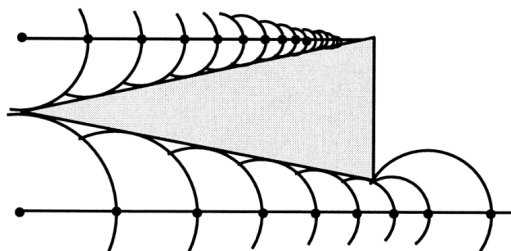


Figure 9: Pseudo code for the geometric implicit surface rendering algorithm

Figure 10: A hit and a miss (image from [3])

The pseudo algorithm for rendering one implicit surface is listed in Figure 9. In the pseudocode D is the maximum ray traversal distance

References

- [1] *Deriving the SDF of a Box*. URL: <https://www.youtube.com/watch?v=62-pRVZuS5c>.
- [2] *Distance Functions*. URL: <https://www.iquilezles.org/www/articles/distfunctions/distfunctions.htm>.
- [3] Jhon C. Hart. “Sphere tracing: a geometric method for antialiased ray tracing of implicit surfaces”. In: *The Visual Computer* (1996).
- [4] *Rendering Implicit Surfaces and Distance Fields: Sphere Tracing*. URL: <https://www.scratchapixel.com/lessons/advanced-rendering/rendering-distance-fields/introduction>.
- [5] Kevin Suffern. *Ray Tracing from the Ground Up*. A K Peters, Ltd., 2016.