



Procedural cloud shader

Project documentation



Field of Studies:	BSc in Computer Science
Specialization:	Computer perception and virtual reality
Author:	Matthias Thomann
Supervisor:	Prof. Urs Künzler
Date:	June 7, 2020
Version:	0.17

Abstract

Clouds contribute a great deal to the overall ambience in games and can be the cherry on top by filling the sky with life. To get as close as possible to real clouds, this project engages in researching and prototyping a procedural, volumetric cloud shader.

In order to achieve volumetric rendering, the document dives into the concept of ray marching, a group of methods used to render a 3D data set inside a container box to make it appear volumetric. Several variants of it are expanded on, like constant step, traditional, and sphere-traced ray marching. Additionally, to account for perception of depth, the volume can be shaded with the aid of surface normal estimation.

In the second part, 2D and 3D noise generation algorithms like Perlin's and the Voronoi algorithm are explained in detail. With fractal Brownian motion, the different layers of noise are then merged into one highly detailed noise texture.

At last, the goal of the project was to create prototypes in Unity Editor displaying both volumetric rendering and noise algorithms, of which all were created successfully. Armed with the combined knowledge of the research results and prototypes, a final shader was created, able to render a completely procedural and volumetric cloudscape.

For future work, the shader could be expanded into a fully-fledged weather system with meteorologically accurate formation of clouds, rain, snow and much more.

Contents

1 General	1
1.1 Purpose	1
1.2 Audience	1
1.3 Revision History	1
2 Natural Clouds	2
2.1 Formation	2
2.2 Types of Clouds	2
2.3 Clouds in Games	3
2.3.1 Skyboxes	3
2.3.2 Billboards	3
2.3.3 Mesh-based Objects	4
2.3.4 Volumetric Clouds	5
3 Volumetric Rendering	6
3.1 Definition	6
3.2 Preliminary Notes	6
3.3 Constant Step Ray Marching	6
3.4 Traditional Ray Marching	8
3.5 Sphere Tracing	9
3.5.1 Signed Distance Functions	9
3.5.2 Sphere Tracing with SDFs	9
3.6 Surface Normals and Lighting	11
3.6.1 Surface Normal Estimation	11
3.7 Shadow Casting	13
3.7.1 Soft Shadows	14
3.8 Shape Blending	15
3.8.1 Constructive Solid Geometry	16
3.9 Ambient Occlusion	17
4 Noise Generation	18
4.1 Random Numbers	18
4.2 2D Random	19
4.3 Procedural Noise Patterns	20
4.3.1 Perlin Noise	20
4.3.2 Voronoi Noise	24
4.3.3 Fractal Brownian Motion	26
5 Project Management	28
5.1 Schedule Adherence	28
5.2 Goal Discrepancies	29
5.3 Future Work	29
5.3.1 Complete Weather System	29
5.3.2 Extensive Lighting Features	29
5.3.3 Measure Realism	29

6 Prototypes and Results	30
6.1 Preliminary Notes	30
6.1.1 Completed Prototypes	30
6.1.2 Dimensions	30
6.1.3 Unity Variables	30
6.2 First Draft	31
6.2.1 Density Sampling	31
6.2.2 Normalizing Density	32
6.3 Improving Noise	33
6.4 Light Transmittance and Light Scattering	34
6.4.1 Sunlight Forward Scattering	34
6.4.2 Directional light	36
6.5 Final Prototype	39
6.5.1 Additional Masking	41
6.6 Realism Check	42
6.6.1 Objective Interpretation	42
6.6.2 Real Life Comparison	43
6.6.3 Convolutional Neural Network	45
6.6.4 Generative Adversarial Network	45
6.6.5 Histogram Comparison	45
6.6.6 Professional Meteorological Assessment	45
6.7 Performance	46
6.7.1 Optimization Attempts	46
Glossary	47
References	49
Listings	51
Figures	51
Code Listings	53

1 General

1.1 Purpose

During this project, all gathered information and knowledge about the researched algorithms and techniques are written down. All prototypes and the final results are documented and compared with real photographs of clouds.

1.2 Audience

This document is written with the intent to further expand existing knowledge about the topic, hence it requires a fundamental knowledge about computer graphics and rendering.

1.3 Revision History

Version	Date	Name	Comment
0.1	March 21, 2020	Matthias Thomann	Initial draft
0.2	March 29, 2020	Matthias Thomann	Added first research results
0.3	April 01, 2020	Matthias Thomann	Added Unity prototype environment
0.4	April 03, 2020	Matthias Thomann	Added further research results
0.5	April 08, 2020	Matthias Thomann	Added further research results
0.6	April 13, 2020	Matthias Thomann	Added further research results
0.7	April 19, 2020	Matthias Thomann	Added research results about noise
0.8	April 26, 2020	Matthias Thomann	Added research results about noise
0.9	May 02, 2020	Matthias Thomann	Added Voronoi noise research
0.10	May 08, 2020	Matthias Thomann	Added FBM noise research
0.11	May 14, 2020	Matthias Thomann	Added prototype results
0.12	May 15, 2020	Matthias Thomann	Added prototype results
0.13	May 19, 2020	Matthias Thomann	Added prototype results
0.14	May 20, 2020	Matthias Thomann	Added prototype results
0.15	June 01, 2020	Matthias Thomann	Added realism checks
0.16	June 03, 2020	Matthias Thomann	Spelling revision
0.17	June 05, 2020	Matthias Thomann	Finalized document

2 Natural Clouds

2.1 Formation

Clouds, as seen in nature, consist of a visible body of tiny water droplets and frozen crystals. In their natural occurrence, clouds are mostly generated from a nearby source of moisture, usually in the form of water vapor. This composition of particles creates the pleasant look of a white-grayish "fluffy" mass, floating in the sky.

Due to certain factors like altitude or water source, different types of cloudscapes can be formed. They vary in shape, convection, density and more. That makes different cloud genera highly unique in terms of appearance.

For now, those factors are regarded as nature's randomness. However, an approximation of randomness will be covered in section 4.

2.2 Types of Clouds

Cloudscapes are classified in multiple groups, mainly differing in altitude, meaning the distance from the earth's surface to the cloud formation. The following four cloud genera stand out due to their distinctiveness. A realistic simulation of a cloud system would consist of a combination of these types, which is why they are displayed here.



Figure 1: Photographic reference of stratus clouds [7].



Figure 2: Photographic reference of cirrus clouds [8].



Figure 3: Photographic reference of an altocumulus cloud formation [9].

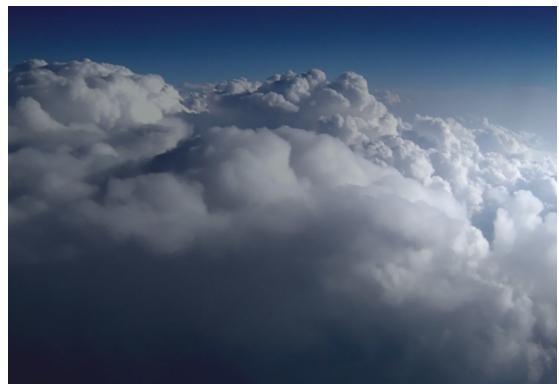


Figure 4: Photographic reference of stratuscumulus cloudscape [10].

2.3 Clouds in Games

Depicted in Figure 3 and Figure 4 of subsection 2.2 are clouds of the genus *cumulus*, which translated to English means *heap* or *pile*. Their remarkable cotton-like look makes them easy to recognize, which is also why they are often used in games as a reference for "normal" clouds. That is also the reason why the prototypes will mainly revolve around this genus. In games, the formation as well as the natural composition are both irrelevant, as the clouds are essentially only used for cinematic ambience or as a medium to enhance the atmosphere. This leaves just the rendering technique and performance to worry about.

2.3.1 Skyboxes

A widespread solution for representing clouds in games is not rendering them separately at all, but instead using a set of polar sky dome images, also known as the skybox. This is a six-sided cube which is rendered around the whole game world. On each inward looking face of the cube, one of the sky dome images is displayed, creating a seamless sky around the inner side of the box.

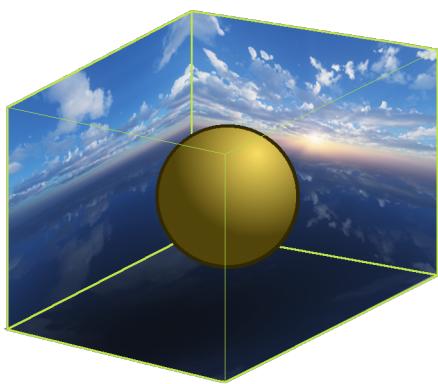


Figure 5: The skybox cube as it is used in games.

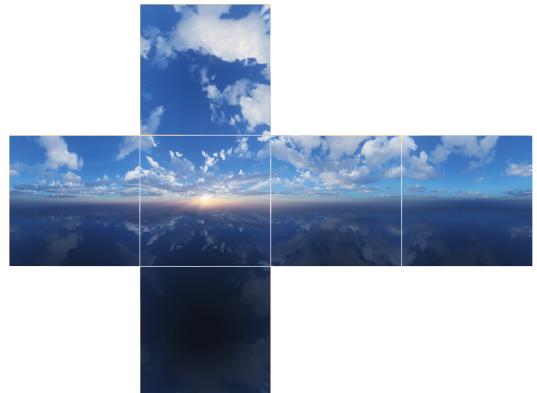


Figure 6: The polar sky dome images, folded out.

Besides rendering the sky, this of course allows clouds to be drawn right into the background. Also, in terms of performance, this is extremely cheap and efficient. On the other hand, it removes the ability for the clouds to move. They also have no volumetric body and no way of interaction with the game world whatsoever.

This method does indeed give the scenery a more cloudy look, but what is missing is the "feel", or in other words the motion, interaction and lifelikeness of the clouds.

2.3.2 Billboards

Similar to the approach with the skybox, this technique also only uses 2D images of clouds. They are rendered individually and are always facing the camera. This is called *billboarding*. Now that each cloud is represented by its own game object, having a position in world space as well as a scale and many other properties, it is possible to animate the clouds. For example, by moving the game objects in a circle around the world, the clouds seemingly "pass by".

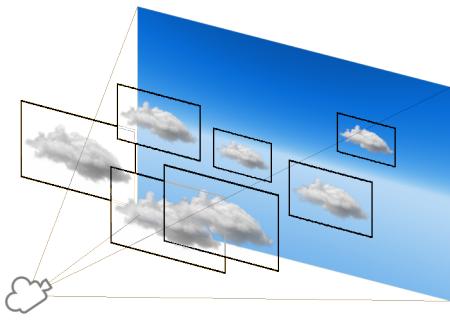


Figure 7: A collection of 2D cloud billboards facing the camera.



Figure 8: The rendered result of the image to the left.

Due to billboarding, the orientation is already given, making the overall time and effort of this technique quite advantageous to others.

The major flaw of using billboards is of course that they are still 2D images, meaning they cannot really change appearance and therefore, do not evolve at all. Still, for many games, this technique suffices in the required diversity of background scenery and does not exceed the allowed performance share for such a task.

2.3.3 Mesh-based Objects

It is imaginable to simply use a polymesh shaped like a cloud and render that like any other game object. By adding a texture, this would make for some decent looking clouds.

However, the level of detail of such a polymesh is directly connected to the amount of vertices and faces that have to be processed every frame. As seen in Figure 9, there are hundreds of polygons required to merely represent the basic shape of a realistic cloud. If a similarly complex mesh is to be used for every cloud, a massive overhead is generated for objects that usually only contribute to the background of a game.

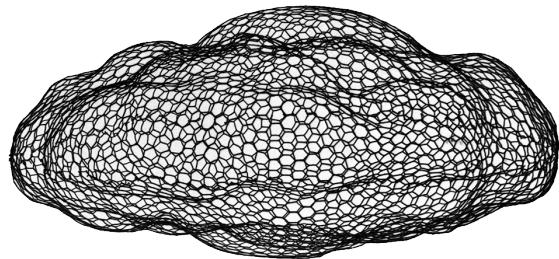


Figure 9: A polymesh in the shape of an altocumulus cloud [11].

Apart from the performance impact, this method offers a volumetric, possibly interactable object just like any other 3D model does. When massively decreasing the polygon count and therefore relinquishing the realistic look, mesh-based objects may be a viable solution for some low poly games. Otherwise, it is not reasonable to use this method.

2.3.4 Volumetric Clouds

Finally, clouds can be rendered via a technique called *volumetric rendering*. The image below shows volumetric cloudscapes as seen in popular AAA titles. The method itself is explained in detail in section 3.

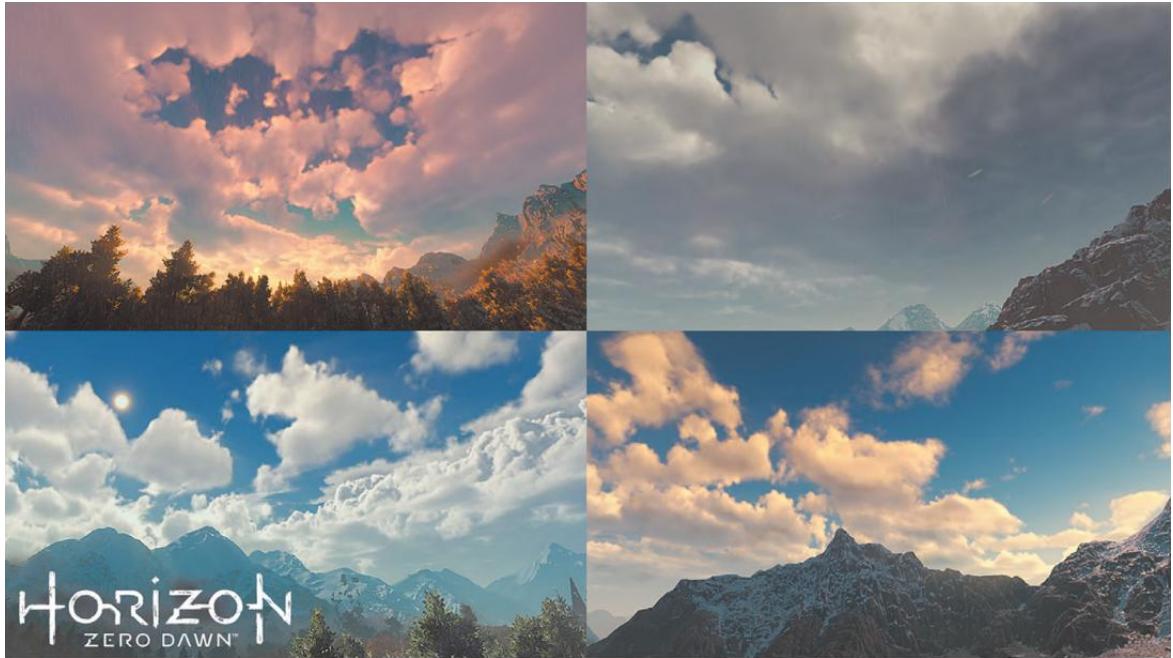


Figure 10: Several volumetric cloudscapes from the game *Horizon: Zero Dawn*, drawn in real time [12].

3 Volumetric Rendering

3.1 Definition

Volumetric rendering describes a technique for generating a visual representation of data that is stored in a 3D volume. This especially comes to use for visual effects that are volumetric in nature, like fluids, clouds, fire, smoke, fog and dust, which are all extremely difficult or even impossible to model with geometric primitives.

In addition to rendering such effects, volumetric rendering has become essential to scientific applications like medical imaging, for which a typical 3D data volume is a set of 2D slice images acquired by a CT (computed tomography) or MRI (magnetic resonance imaging) scanner.

The data volume is also called a *scalar field* or *vector field*, which associates a scalar or vector value, called *voxel* (short for *volume element*), to every point in the defined space. For a scalar field, it can be imagined like a 3D grid, where each point holds a single number. This number could, for example, represent the density of a cloud at that very point.

3.2 Preliminary Notes

Most of the figures in the upcoming subsections depict only a single ray. However, this is only for explanatory purposes. The process has to be executed not once, but for each fragment processed by the fragment shader.

3.3 Constant Step Ray Marching

To actually render the volume data, a method called *ray marching* is used. With it, the surface distance of the volumetric data is approximated by creating a ray from the camera to the object for each fragment. The ray is then extended into the volume of the object and stepped forward until the surface is reached.

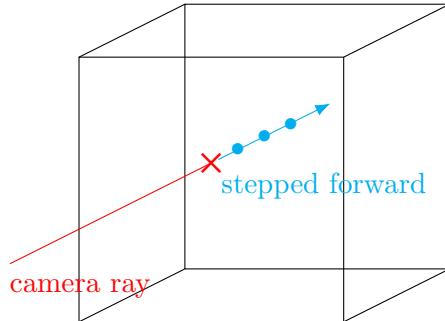


Figure 11: Ray marching concept visualized.

The ray-surface intersection is not directly calculated because it is too complex for volumes like clouds, which is why the surface distance is approximated instead.

In ray marching, the algorithm only knows when it has reached the surface, or to be precise, when it is inside the actual object volume.

With this information, it is only possible to extend the ray in steps of a predefined length until the inside of the object is reached. With a constant step, the approximation of the surface distance is exactly as precise as the size of the constant step.

Once the ray is inside the actual volume, the function returns the distance for this ray.

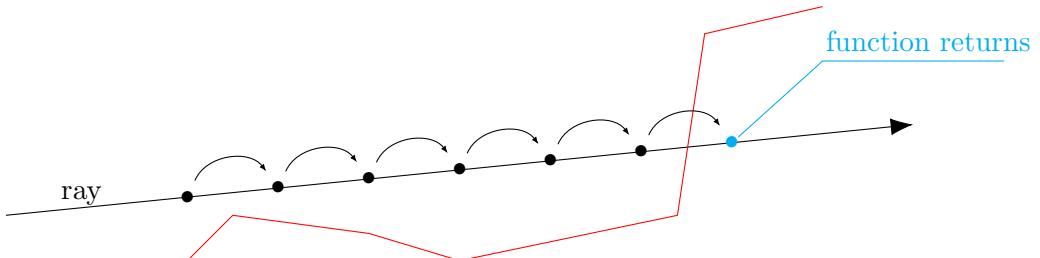


Figure 12: Traditional ray marching.

An implementation of this algorithm can be seen in Listing 1. Note that the volume to be rendered in this example is just a simple sphere.

```

1 fixed4 raymarch(float3 position, float3 direction)
2 {
3     for (int i = 0; i < MAX_STEPS; i++)
4     {
5         if (sphereHit(position))
6             return fixed4(1,0,0,1);
7
8         position += normalize(direction) * STEP_SIZE;
9     }
10
11    return fixed4(0,0,0,1);
12 }
```

Listing 1: Implementation of a ray march function with constant step.

In order to check if the ray is inside the volume, the function `sphereHit()` is used.

```

1 bool sphereHit(float3 position) {
2     float4 sphere = float4(0, 1, 0, 1);
3     return distance(sphere.xyz, position) < sphere.w;
4 }
```

Listing 2: Implementation of a volume distance function for a sphere.

3.4 Traditional Ray Marching

It is obvious to see that for a constant step ray march to result in an accurate approximation of the surface distance, the step size is required to be relatively small. This has a direct impact on performance and thus, is not a viable solution for the problem.

In traditional ray marching, an optimization for that has been developed. The algorithm does not blindly step forward, but instead tries to get as close to the real distance as possible. After the volume is reached, the step size is decreased and the ray steps out of the volume again. It then tries to approximate the surface distance by stepping back and forth repeatedly in continuously smaller steps, thus converging towards the exact intersection. Once the step size falls below a certain threshold, the distance approximation is assumed to be precise enough and the value is returned for that ray march.

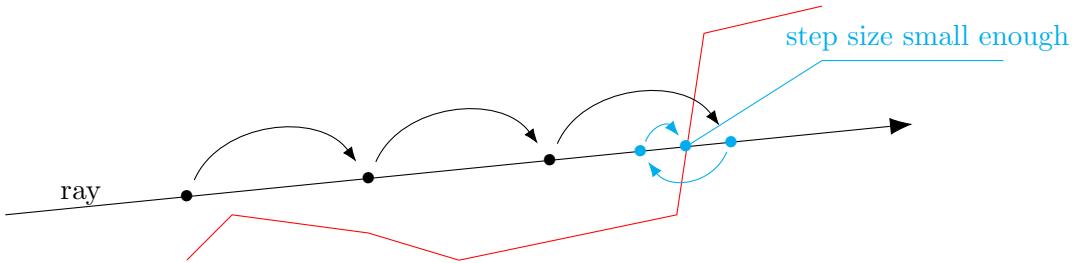


Figure 13: Traditional ray marching.

As clearly visible, the traditional ray marching ends up with a more accurate result and the amount of steps per ray could be relatively lower, ultimately saving performance.

However, there is still an issue. The algorithm may jump in and out of the volume, even if it would already be precise enough, essentially taking unnecessary steps.

```

1 fixed4 raymarch(float3 position, float3 direction)
2 {
3     float stepSize = STEP_SIZE;
4     float dirMultiplier = 1;
5     for (int i = 0; i < MAX_STEPS; i++)
6     {
7         if (stepSize < MINIMUM_STEP_SIZE)
8             return fixed4(1,0,0,1);
9
10        if (sphereHit(position)) {
11            // reduce step size by half and invert marching direction.
12            stepSize /= 2;
13            dirMultiplier = -1;
14        } else {
15            dirMultiplier = 1;
16        }
17
18        position += normalize(direction) * stepSize * dirMultiplier;
19    }
20
21    return fixed4(0,0,0,1);
22 }
```

Listing 3: Implementation of a traditional ray march function with converging surface distance approximation.

3.5 Sphere Tracing

An even better approach to approximate the intersection of the ray and the volume is called *sphere tracing*. Instead of evaluating if the ray is inside the volume or not, an exact distance to the scene is measured. This distance is the minimum amount of space the algorithm can march along its ray without colliding with anything. For that, a function group called *signed distance functions* is used.

3.5.1 Signed Distance Functions

A signed distance function (SDF) returns the shortest distance from a given point in space to some surface. The sign of the returned value indicates whether that point is inside the surface or outside, hence the name.

For example, the signed distance function $f(p)$ for a point $p = (p_1, p_2, p_3)$ to the surface of a sphere $s = (s_1, s_2, s_3)$ with radius R looks like this:

$$f(p) = \sqrt{(s_1 - p_1)^2 + (s_2 - p_2)^2 + (s_3 - p_3)^2} - R$$

This translates into a simple code snippet, mostly identical to the function `sphereHit()` in Listing 2, except the distance is returned instead of a Boolean.

```
1 float sceneSDF(float3 position) {
2     float4 sphere = float4(0, 0, 0, 1);
3     return distance(sphere.xyz, position) - sphere.w;
4 }
```

Listing 4: Implementation of a signed distance function for a sphere.

With the sphere in the example being at the origin and having $R = 1$, a positive distance is returned for points outside the sphere and a negative distance if the point is inside the sphere.

```
1 float d1 = sceneSDF(float3(2, 0, 0)); // d1 = 1.0
2 float d2 = sceneSDF(float3(0, 0.5, 0)); // d2 = -0.5
3 float d3 = sceneSDF(float3(5, -5, 5)); // d3 = 7.66
```

3.5.2 Sphere Tracing with SDFs

If the distance to the scene can be calculated with a signed distance function, the algorithm becomes rather straight forward. The distance to the scene is evaluated at the start, then one can freely march along the ray for that amount of distance. Once arrived at the new point, the process is repeated until the SDF returns a small enough value.

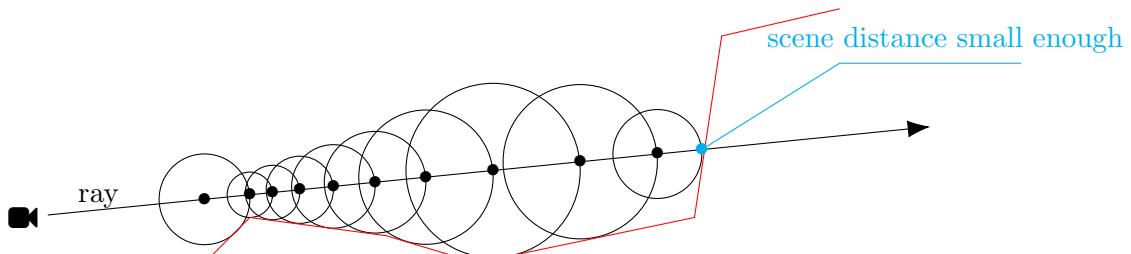


Figure 14: Ray marching with SDF-based sphere tracing.

As seen in Figure 14, the result is highly accurate. For the previous example with just one single sphere as a volume, the algorithm can be implemented like in Listing 5.

```

1 float raymarch (float3 position, float3 direction)
2 {
3     float dOrigin = 0.0;
4     for (int i = 0; i < MAX_STEPS; i++)
5     {
6         float dScene = sceneSDF(position + dOrigin * direction);
7         if (dScene < SURFACE_DISTANCE || dScene > MAX_DISTANCE)
8             break;
9
10        dOrigin += dScene;
11    }
12    return dOrigin;
13 }
```

Listing 5: Implementation of ray marching with sphere tracing.

In order to save on performance, it is imperative to break the loop when `distanceScene` exceeds `MAX_DISTANCE`. This way, the distance evaluation for that ray can be stopped earlier than waiting for the loop to complete. Another example why this check is important can be seen in the next figure. The ray is terminated early, because it does not collide and never reaches the minimum surface distance.

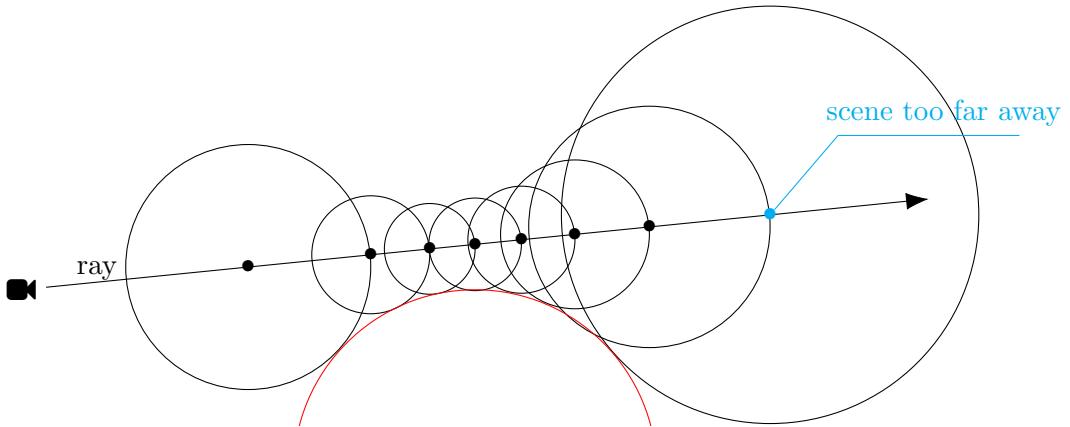


Figure 15: Ray marching with SDF-based sphere tracing, without collision.

3.6 Surface Normals and Lighting

As it is the case for many other lighting models, the surface normals are used to calculate lighting in volumetric rendering. If the object is defined with a polymesh, the surface normals are usually specified for each vertex. The normals for any given point on the surface can then be calculated by interpolating the adjacent vertex normals.

Since there is no polymesh in volumetric rendering, another solution has to be found for calculating the surface normals for a scene defined by signed distance functions. Because of that, it is not possible to explicitly calculate the normals and therefore, an approximation is used.

3.6.1 Surface Normal Estimation

To approximate the normal vectors in a 3D data volume, the *gradient* is used. The gradient represents the direction of greatest change of a scalar function. In Figure 16, the red arrows visualize the gradient. It is similar in 3D, where the gradient can be described as the path a ball would follow rolling downwards when dropped from the top of a corner.

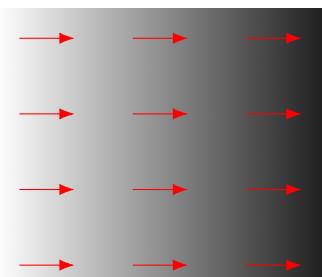


Figure 16: Gradient in a 2D scalar field.

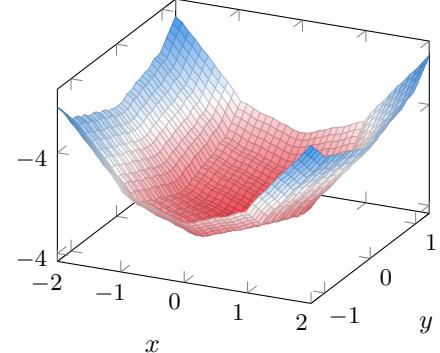


Figure 17: Gradient in a 3D scalar field.

Mathematically, the gradient of a function f at point $p = (x, y, z)$ defines the direction to move in from p to most rapidly increase the value of f . It is written as ∇f .

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$$

Instead of calculating the real derivative of the SDF, an approximation is used to estimate the normal vectors. As previously declared, the signed distance function returns zero for a point on the surface, greater than zero if the point is outside and less than zero if it is inside the volume. Therefore, the direction at the surface which will go from negative to positive most quickly will be orthogonal to the surface.

The estimation \vec{n} is done by sampling some points around the point on the surface and take their difference, the result of which is the approximate surface normal.

$$\vec{n} = \begin{bmatrix} f(x + \epsilon, y, z) - f(x - \epsilon, y, z) \\ f(x, y + \epsilon, z) - f(x, y - \epsilon, z) \\ f(x, y, z + \epsilon) - f(x, y, z - \epsilon) \end{bmatrix}$$

The implementation of surface normal estimation looks like this:

```

1 float3 estimateNormal(float3 p) {
2     return normalize(float3(
3         sceneSDF(p + float3(EPSILON, 0, 0)) - sceneSDF(p - float3(EPSILON, 0, 0)),
4         sceneSDF(p + float3(0, EPSILON, 0)) - sceneSDF(p - float3(0, EPSILON, 0)),
5         sceneSDF(p + float3(0, 0, EPSILON)) - sceneSDF(p - float3(0, 0, EPSILON)),
6     ));
7 }
```

Listing 6: Implementation of surface normal estimation.

Now that the normal vectors can be calculated for the volume, the object can be shaded. In this example, the Phong Illumination Model [13] is used.

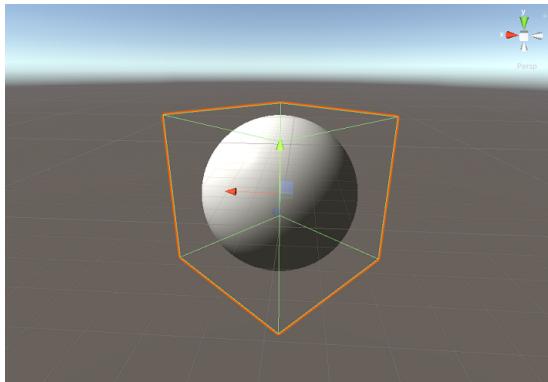


Figure 18: A 3D cube with a volumetric shader.

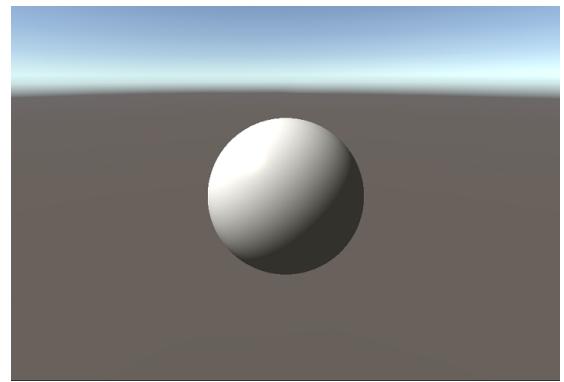


Figure 19: The shaded sphere rendered volumetrically.

3.7 Shadow Casting

In ray marching, rendering cast shadows proves to be rather easy. Naturally, the light ray comes from the sun, bounces off in the world and may eventually hit the eye of the observer. Since only a minute fraction of those rays actually reach the observer (the camera), a huge amount of rays would be calculated for nothing. Consequently, the rays are not traced from the light source to the camera but the other way around instead.

As defined in Listing 5, the `raymarch()` function moves along the given ray and returns the distance to the intersection point of ray and volume. Therefore, when a surface point has been determined, a second ray march can be started from the newly found point in the opposite direction of the primary light source's direction. If anything is hit on the way, the surface point lies in the shadow of the second hit object and should be darkened.

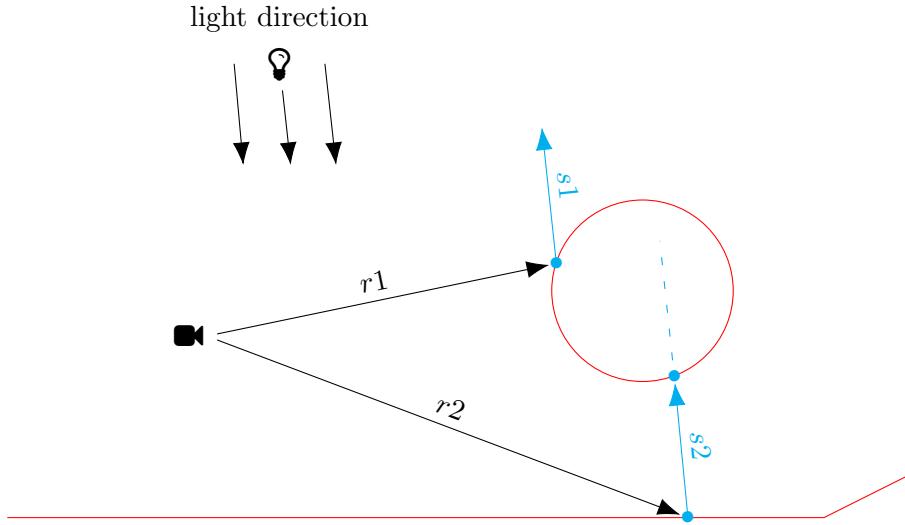


Figure 20: Shadow casting in ray marching.

As seen in the figure above, the ray r_1 hits the volumetric sphere, then checks if anything is between the ray intersection and the negative light source direction. In this case, s_1 does not collide with anything and the surface is shaded normally. For the other ray r_2 however, the shadow ray march returns a distance $s_2 > 0$ and $s_2 < \text{MAX_DISTANCE}$, meaning some object is in-between the hit point and the light source, casting a shadow.

```

1 float hardshadow(float3 position, float3 direction, float dMin, float dMax)
2 {
3     float dOrigin = dMin;
4     for (int i = 0; i < MAX_STEPS; i++) {
5         float dScene = sceneSDF(position + direction * dOrigin);
6         if (dScene < SURFACE_DISTANCE)
7             return 0.0;
8         if (dScene > dMax)
9             return 1.0;
10
11         dOrigin += dScene;
12     }
13     return 1.0;
14 }
```

Listing 7: Implementation of hard shadow casting.

It is very clearly similar to SDF-based sphere tracing, except that only 0 or 1 is returned instead of the distance. The final color is then multiplied by this output. For 0, this results in a total black, hence the name *hard* shadows.

3.7.1 Soft Shadows

The method described in Figure 20 evaluates only if any given point is directly covered by any other object. It does not account for diffuse shadows with soft edges, called *penumbra* or simply *soft* shadows. But there is an easy and also cost-effective solution to that problem. Instead of strictly returning 0 when an object is covered by another, the shortest distance to scene (qualified by some factor k) is returned.

```

1 float softshadow(float3 position, float3 direction, float dMin, float dMax,
                  float k)
2 {
3     float result = 1.0;
4     float dOrigin = dMin;
5     for (int i = 0; i < MAX_STEPS; i++) {
6         float dScene = sceneSDF(position + direction * dOrigin);
7         if (dScene < SURFACE_DISTANCE)
8             return 0;
9         if (dOrigin > dMax)
10            return result;
11
12         result = min(result, k * dScene / dOrigin);
13         dOrigin += dScene;
14     }
15     return result;
16 }
```

Listing 8: Implementation of hard shadow casting.

Those are the resulting renders with a sphere and a flat box as the volumetric scene.

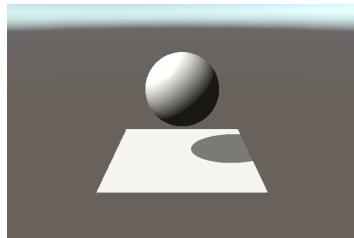


Figure 21: Hard shadows only.

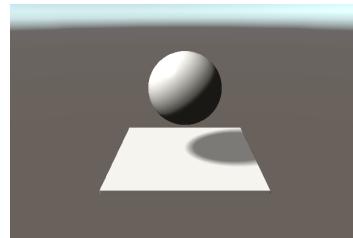


Figure 22: Soft shadows with $k = 7.0$.

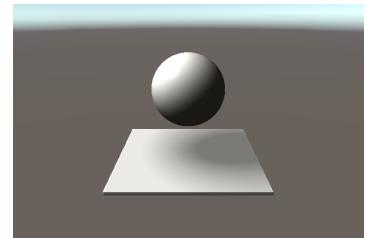


Figure 23: Soft shadows with $k = 1.2$.

3.8 Shape Blending

Another thing that comes free with ray marching is *shape blending*. It describes the concept of blending the signed distance functions of multiple shapes together with this simple method:

```
1 float blend(float d1, float3 d2, float k)
2 {
3     return k * d1 + (1 - k) * d2;
4 }
```

Now two shapes can simply be blended like that:

```
1 float sceneSDF(float3 position)
2 {
3     return blend(sphereSDF(position), boxSDF(position), 0.5);
4 }
```

The following image displays the two blended shapes. Due to the fact that the shadow calculation is based on the same SDFs, no additional changes have to be made in this regard.

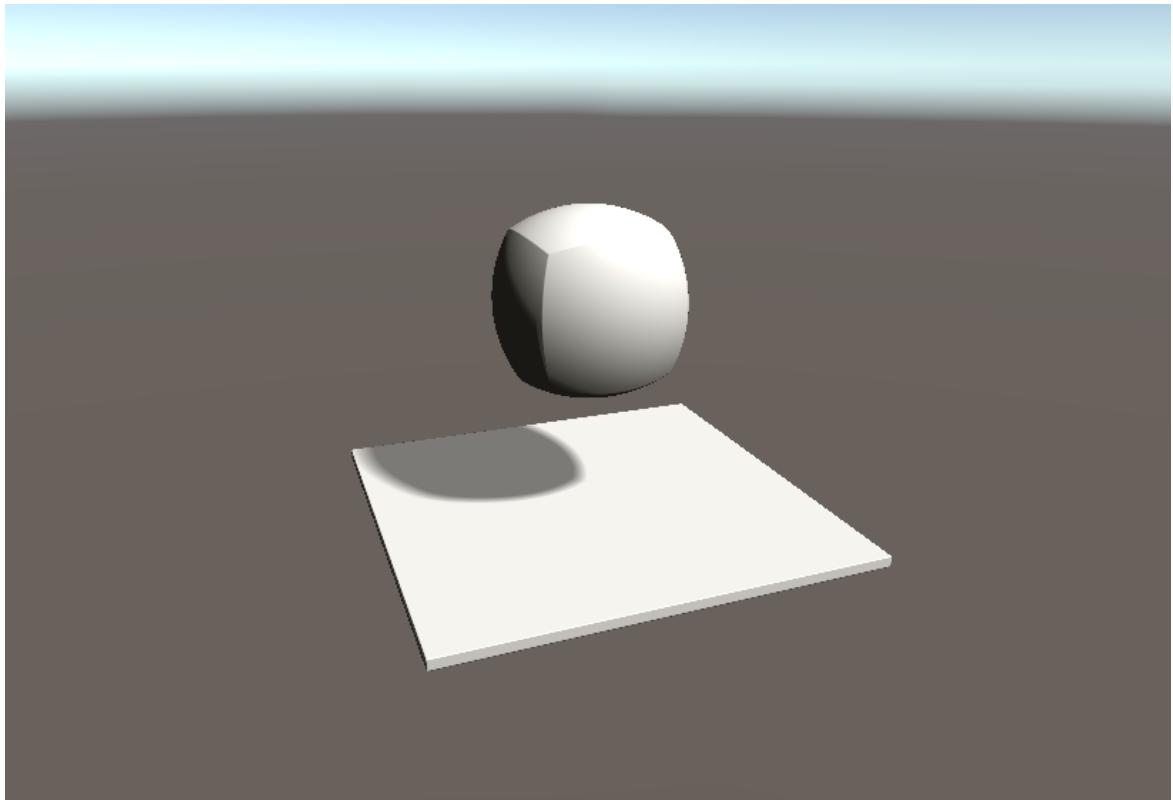


Figure 24: A blended sphere and box SDF with $k = 0.5$.

3.8.1 Constructive Solid Geometry

To create more interesting figures than a rounded box, constructive solid geometry (CSG) operators can be used. As seen in Figure 25, "holes" are cut into the geometry. This is done by taking the difference (or intersection) of the box and a cylinder that goes through the box. Like the `blend()` function takes in two signed distance function results, the following methods are also based on those values.

```
1 float intersection(float d1, float d2)
2 {
3     return max(d1, d2);
4 }
5
6 float union(float d1, float d2)
7 {
8     return min(d1, d2);
9 }
10
11 float difference(float d1, float d2)
12 {
13     return max(d1, -d2);
14 }
```

Listing 9: Implementation of constructive sold geometry.

In this example, the intersection was done three times, for each axis once.

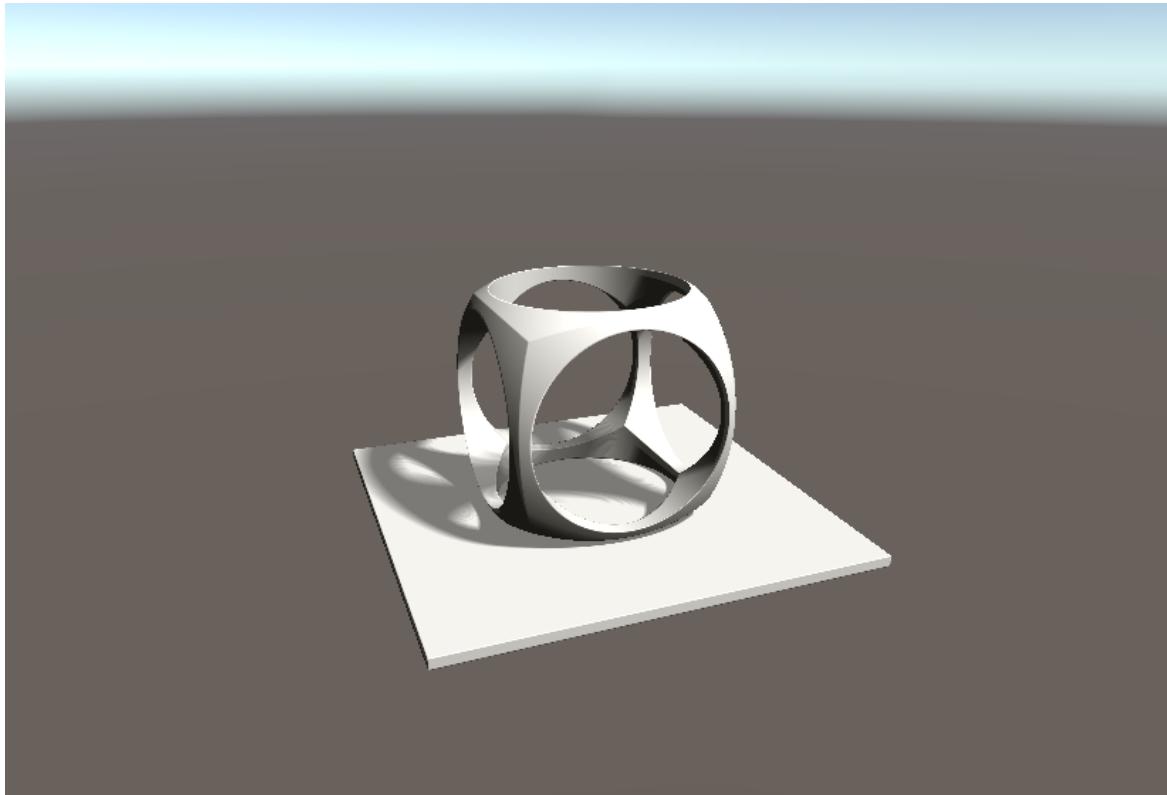


Figure 25: A blended sphere and box with holes along each axis.

3.9 Ambient Occlusion

Shadow casting already looks quite realistic, but there is an important detail missing, called *ambient occlusion*. This method darkens areas around edges and crevices in the scene, making them look less exposed to the light and its environment. The algorithm for that is fairly uncomplicated and straightforward, given all the previously defined methods like `sceneSDF()` and `raymarch()` already exist.

When the `raymarch()` function returns a valid distance, a surface is hit. On that hit point p_1 , the normal vector \vec{n} is estimated. Now the distance to the nearest surface in the direction of \vec{n} is evaluated. If on that ray a hit point p_2 is close, the color for the original hit point p_1 is darkened by some amount, depending on how far apart those points are.

```

1 float ambientOcclusion(float3 p, float3 direction) {
2     float ao = 0;
3     float dOrigin = 0;
4
5     for (int i = 1; i <= AO_ITERATIONS; i++) {
6         dOrigin = AO_STEP_SIZE * i;
7         ao += max(0, dOrigin - sceneSDF(p + direction * dOrigin)) / dOrigin;
8     }
9     return 1 - ao * AO_INTENSITY;
10 }
```

Listing 10: Implementation of ambient occlusion.

This comes close to the constant step ray marching algorithm, since it is marched along the ray in a predefined step size. On line 7, the scene SDF is subtracted from the total distance and then divided by it. This just puts the scene distance in relation to the total distance. Also, `max()` is used because the SDF can return a negative number for points inside the surface, so in order to not brighten the scene at point p when this is the case, 0 is used instead.

With `AO_STEP_SIZE = 0.1`, `AO_ITERATIONS = 3` and `AO_INTENSITY = 0.2`, the following output is produced:

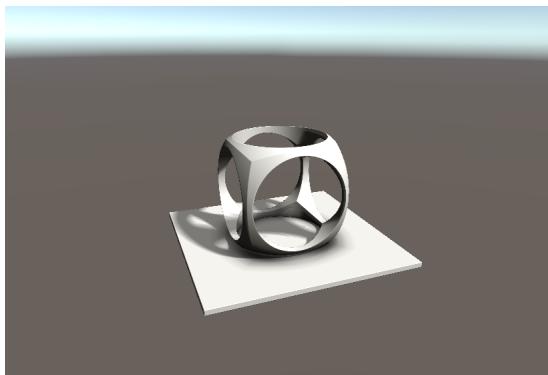


Figure 26: Ambient occlusion applied to the scene.

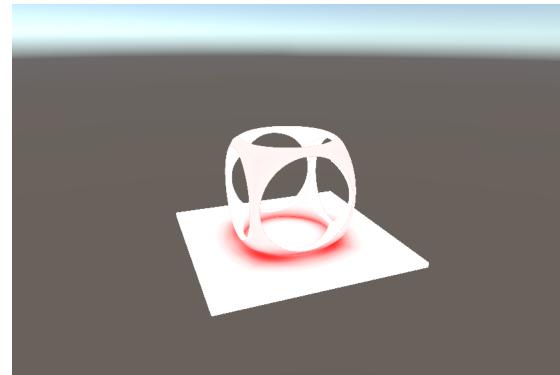


Figure 27: Only the ambient occlusion part drawn in red.

When comparing the previous Figure 25 with Figure 26, the darker ground around the object clearly improves the scene.

4 Noise Generation

Nature's unpredictability plays a big role in the diversity and appearance of cloudscapes. In shaders, an approach to that *randomness* is used called *noise generation*. In order to be able to implement random noise generation, several important topics need to be looked into. It is best to start with randomness in computer science and how it is handled inside a shader program.

4.1 Random Numbers

Unfortunately, there is no magic function which returns a pure random number inside the seemingly predictable and rigid code environment. So the question arises as to how such randomness can be generated.

For this, the function $rnd(x) = fract(\sin(x))$ is inspected, where $fract(x) = x - floor(x)$.

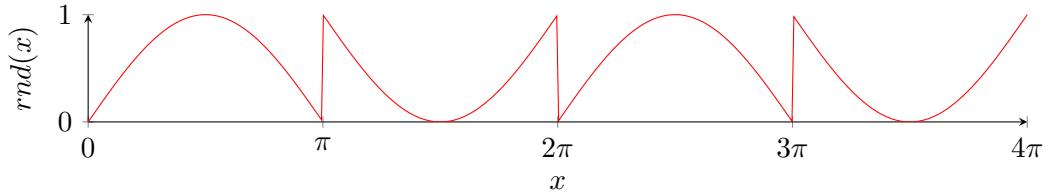


Figure 28: Random numbers with the fractional value of sine of x.

The sine values fluctuate between -1.0 and 1.0 , but with *fract*, only the fractional part is evaluated, turning the negative values into positive ones. This effect can be used to get some pseudo-random values by "compressing" the function horizontally, or in other words by increasing the frequency of the sine wave.

The next figure displays the function $rnd(x) = fract(\sin(x) * 10000)$.

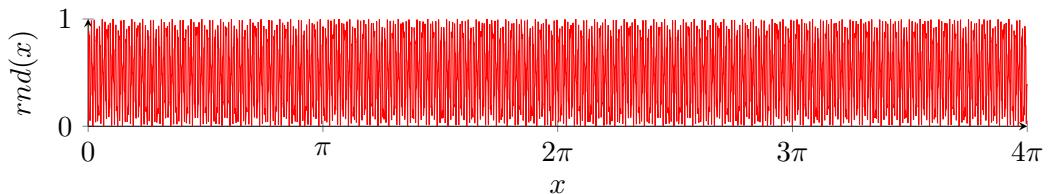


Figure 29: Random numbers with the fractional value of sine of x multiplied by 10000.

It is clearly visible that the function $rnd(x)$ became chaotic and returns practically random values. However, it is noteworthy that $rnd(x)$ is still a deterministic function, which means for example $rnd(1.0)$ is always going to return the same value.

4.2 2D Random

To generate a pseudo-random number from two and three values instead of one, the same function can be used, with some tweaks. Those two numbers come as a two-dimensional vector, which needs to be transformed into a single floating point number. According to Vivo, the dot product is particularly helpful in that case [5]. It returns a single float value between 0.0 and 1.0 depending on the alignment of two vectors. They describe the following methods:

```
1 float random(float2 co) {
2     float2 other = float2(12.9898, 78.233);
3     return fract(sin(dot(co, other)) * 43758.5453123);
4 }
```

Listing 11: Implementation of 2D random number generation.

```
1 float random(float3 co) {
2     float3 other = float3(12.9898, 78.233, 37.719);
3     return fract(sin(dot(co, other)) * 43758.5453123);
4 }
```

Listing 12: Implementation of 3D random number generation.

When using the fragment coordinates as the vector `co` to call `random(co)` for every pixels, the resulting image shows a seemingly random assortment of pixels holding values from 0 to 1 (from black to white).

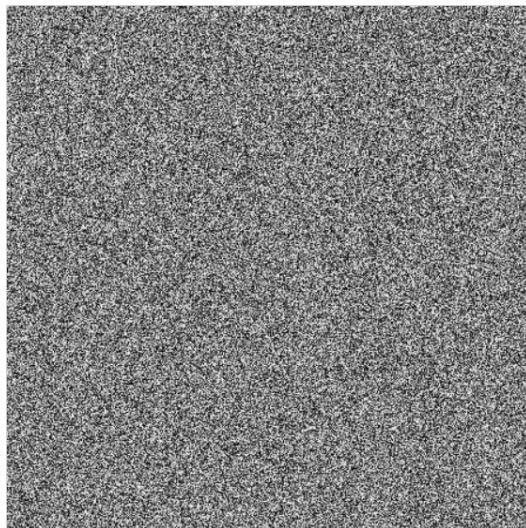


Figure 30: 2D random function visualized.

This method of procedural randomness still has one major flaw: It has no patterns. Contradictory to the word *random*, a certain pattern is required in order to generate *random* clouds. Luckily, there is more to random generation than just a highly sped up sine wave.

4.3 Procedural Noise Patterns

Now that the concept of random numbers in the world of shaders is no longer a mystery, more advanced noise generation algorithms can be introduced. When using the word *noise* in this context, usually procedural pattern generation is meant.

4.3.1 Perlin Noise

One of the most commonly used procedural pattern generation algorithms is that of Ken Perlin. Named after him, the algorithm works with the gradient, which was already introduced in subsection 3.6.1.

It consists of the following three steps:

1. grid definition
2. dot product calculation between random gradient and distance vectors
3. interpolation of those dot product values

Note that the following example refers to two-dimensional Perlin noise generation, but with some tweaks, is very much applicable for higher dimensional noise generation.

First, the 2D image space is split into a grid. For each vertex or corner point on this grid, a pseudo-random gradient vector is determined.

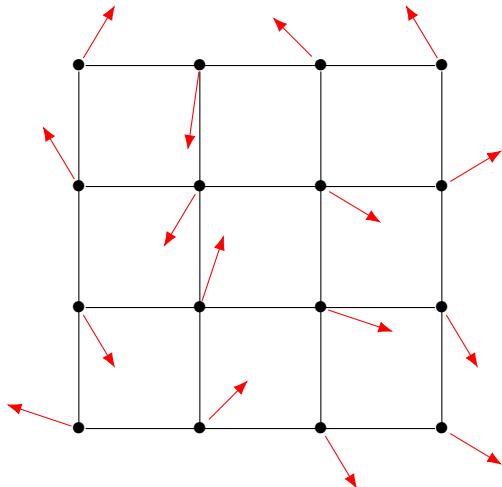


Figure 31: Perlin grid with pseudo-random gradient vectors.

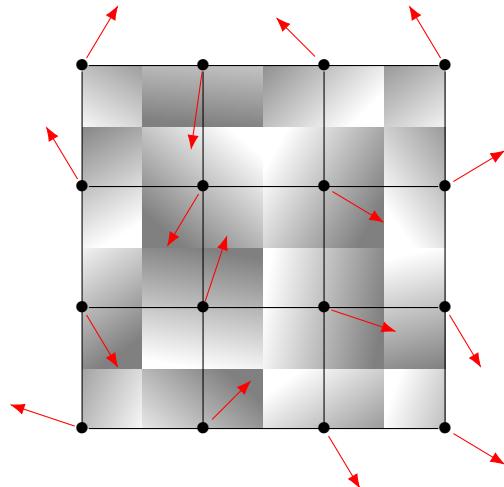


Figure 32: Perlin grid with visualized gradient vectors.

For the next step, it is easier to only inspect a single cell. Given the algorithm currently processes the highlighted pixel p in Figure 33, the next task is to determine the distance vectors from each adjacent corner point to the that pixel. Note that in \mathbb{R}^2 , the amount of corners is four, while in \mathbb{R}^3 , its eight.

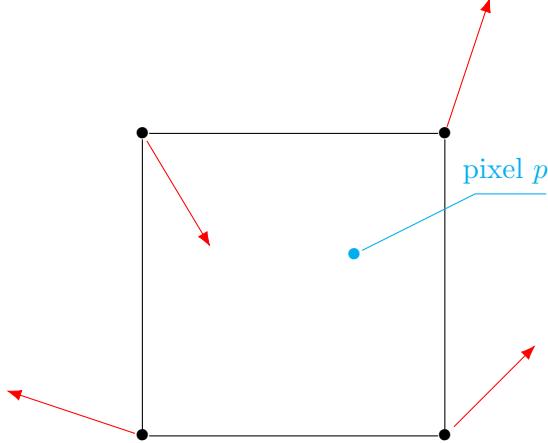


Figure 33: Perlin grid cell with gradient vectors.

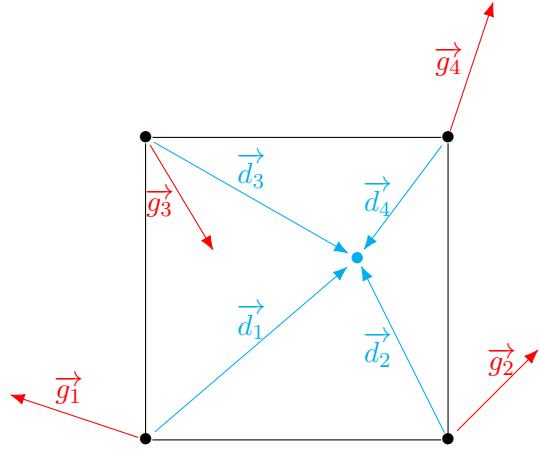


Figure 34: Perlin grid cell with distance vectors from each vertex to the pixel.

Then, the dot product is calculated for each distance vector and its gradient vector. This qualifies how similar those two vectors are, returning a positive number if they face the same direction and a negative one for the opposite. The dot product is 0 if the vectors are perpendicular.

$$\begin{aligned}s &= g_1 * d_1, \\t &= g_2 * d_2, \\u &= g_3 * d_3, \\v &= g_4 * d_4.\end{aligned}$$

The values s, t, u, v represent the influences of the respective gradient on the final color of the pixel p . When visualizing those values as vectors with their length being the influence, it looks like this:

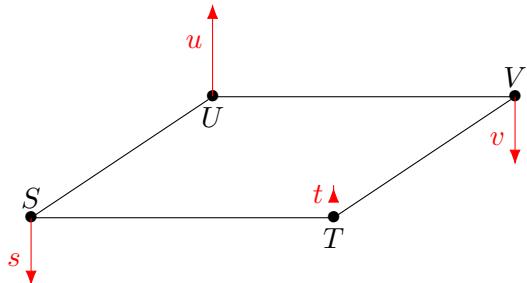


Figure 35: Perlin grid cell with visualized influences of gradient vectors.

It is clearly recognizable that the color of the pixel is influenced the most by v . Now those four numbers can be combined into one final number, the color value. For that, some sort of average calculation is used. For \mathbb{R}^2 , the following ruleset applies:

1. find the average of the first pair of numbers
2. find the average of the second pair of numbers
3. average those two numbers together

To get an accurate mean value of those influences, rather than using the arithmetic average, a weighted average calculation is used. The weight for that is how close p is to the vertices. This means if p is close to a corner point, the influence of that vertex should be weighted heavier than the influences of all other corner points.

This is solved by linear interpolation.

$$d_x = (T_x - p_x)/(T_x) - (S_x), \\ d_y = (U_y - p_y)/(U_y) - (S_y).$$

$$w_1 = \text{lerp}(u, v, d_x), \\ w_2 = \text{lerp}(s, t, d_x), \\ w_{\text{final}} = \text{lerp}(w_1, w_2, d_y).$$

Both variables d_x and d_y represent the interpolation weight, being between 0 and 1. With w_1 , the interpolation between s and t is done, depending on how far to the right the pixel is, related to its cell. This results in the first interpolation of the X-axis. Now w_2 is calculated, giving the second horizontal value in-between u and v . Finally, both w_1 and w_2 are linearly interpolated in relation to d_y , which gives the final average number.

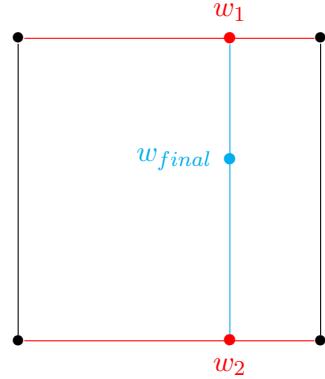


Figure 36: Perlin vertex weights in 2D space with four corners and three interpolations.

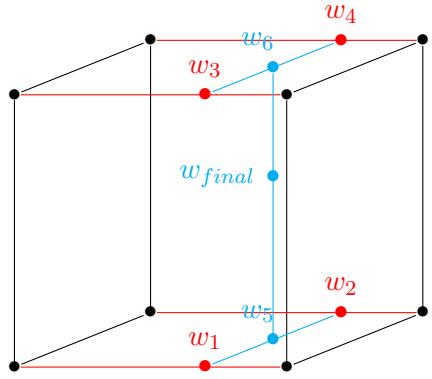


Figure 37: Perlin vertex weights in 3D space with eight corners and seven interpolations.

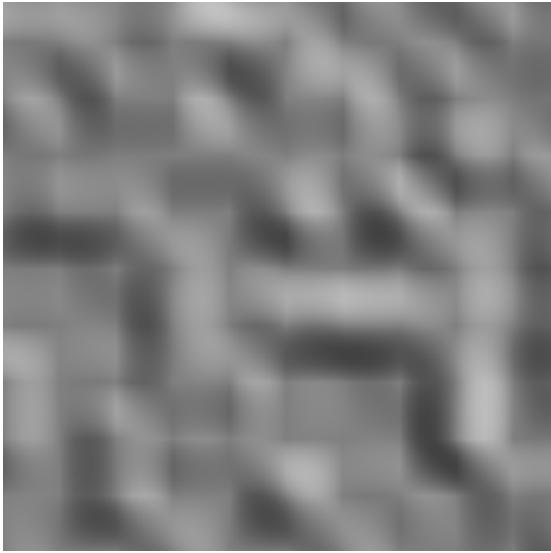


Figure 38: 2D Perlin noise texture with a 10x10 grid.

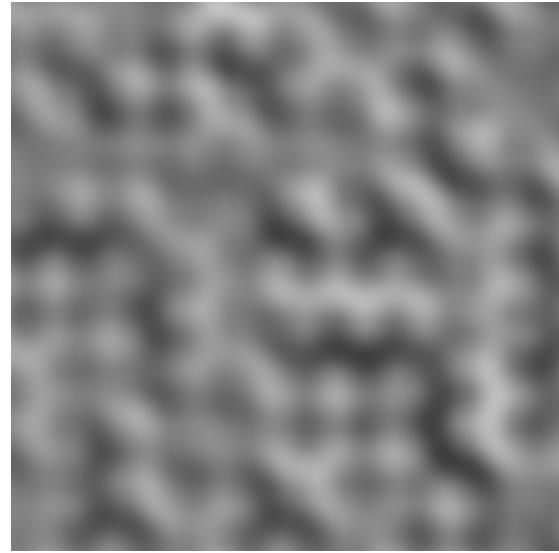


Figure 39: 2D Perlin noise texture with Perlin's fade function.

By default, the Perlin noise texture shows a significant amount of artifacts along the grid lines. This can be fixed by using Perlin's fade function [14] for d_x and d_y , which is defined by $f(t) = 6t^5 - 15t^4 + 10t^3$.

For 3D, Perlin describes that rather than calculating random gradient vectors, a simple set of 12 distinct vectors can be used, which still provides sufficient randomness but is faster [15]. For each grid corner, a hash function is used to generate an index (from 0 to 11), with which one of the gradient vectors is then chosen.

4.3.2 Voronoi Noise

While Perlin's noise algorithm is heavy on vector calculation and interpolation, other noise patterns are less complex to understand and construct, like the *Voronoi* noise, also known as *Worley* or *cellular* noise. The name derives from its similar structure to a Voronoi diagram, in which points, called *seeds*, are randomly scattered inside a defined space. After that, regions are created, consisting of all points closer to that seed than to any other.

As for the noise pattern, there are some alterations. To get a more even distribution, the noise algorithm starts by dividing the space into a grid, for which each cell is assigned a random point. From there, each fragment gets colored by how far it is to the closest seed in its cell.

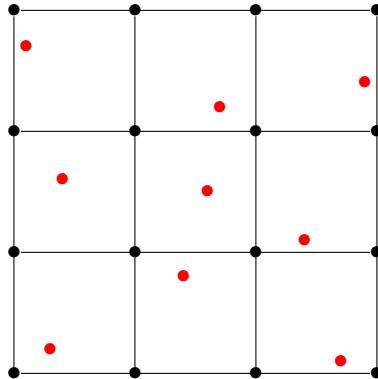


Figure 40: Voronoi grid with pseudo-randomly assigned seed points for each cell.

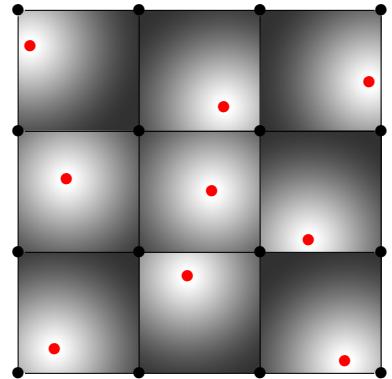


Figure 41: Voronoi grid with seed distances visualized.

As understandable, in Figure 41, hard contours are still visible along the grid lines. The final step is done by including the adjacent cells when finding the closest seed for any given fragment. This amounts to $3^n - 1$ neighboring cells, where n is the number of dimensions. This means for 2D space its eight cells, while in 3D its 26.

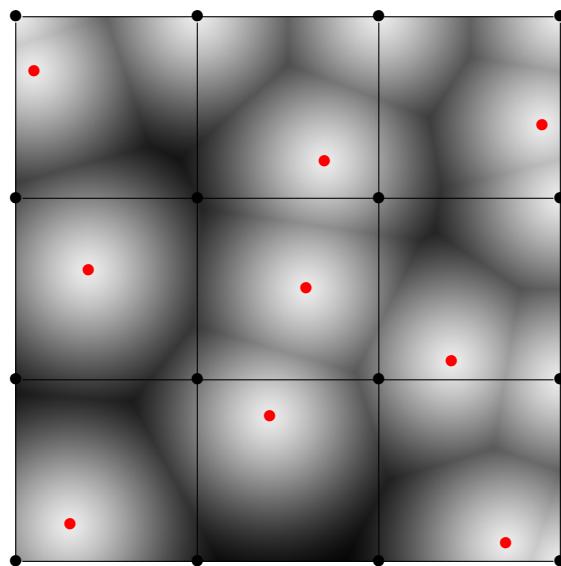


Figure 42: Complete 2D Voronoi noise pattern.

An implementation of this relatively simple algorithm could look like the following listing. In it, `randomSeed()` is used like the previously introduced function `random()`, except that it returns a two-dimensional vector instead of a scalar. With that, a deterministically random point can be generated for any given cell.

```

1 float2 randomSeed(float2 co) {
2     return float2(
3         fract(sin(dot(co, float2(12.9898, 78.233))) * 43758.5453123),
4         fract(sin(dot(co, float2(39.3461, 11.135))) * 14375.8545359));
5 }
6
7 float voronoi(float2 p) {
8     float2 pCell = floor(p);
9     float dMin = 999;
10
11    for(int x = -1; x <= 1; x++) {
12        for(int y = -1; y <= 1; y++) {
13            float2 cell = pCell + float2(x, y);
14            float2 seed = cell + randomSeed(cell);
15            float d = distance(seed, p);
16            if (d < dMin) {
17                dMin = d;
18            }
19        }
20    }
21
22    return dMin;
23 }
```

Listing 13: Implementation of 2D Voronoi noise algorithm.

Since the Voronoi noise algorithm creates a cellular pattern, it is well suitable for simulating natural distribution of cloud heaps, as they are in some way also formed "in cells".

4.3.3 Fractal Brownian Motion

In the world of shaders, the term *fractal Brownian motion* (fBm) is often described as adding different levels of noise together, thus creating a self-similar pattern across different scales [16]. This simplified description meets the required level of detail for this section, a complete explanation and derivation of the fractal Brownian motion is beyond the scope of this paper.

In shaders, fBms are also called *fractal noise*. They are usually implemented by adding different iterations of noise (called *octaves*), while successively incrementing the frequencies in regular steps (*lacunarity*) and decreasing the amplitude (*gain*) of the noise. This results in a more detailed noise, meaning a finer granularity of the pattern in the noise.

```

1 #define LACUNARITY 2.0
2 #define GAIN 0.5
3 #define OCTAVES 1
4
5 float fbm(float2 p) {
6     float frequency = 1.0;
7     float amplitude = 0.5;
8
9     float total = 0;
10    float maxValue = 0;
11    for(int i = 0; i < OCTAVES; i++) {
12        float current = noise(p * frequency) * amplitude;
13        total += current;
14        maxValue += amplitude;
15
16        amplitude *= GAIN;
17        frequency *= LACUNARITY;
18    }
19
20    return total/maxValue;
21 }
```

Listing 14: Implementation of fractal Brownian motion function.

Interestingly, the only things that change for 3D is `float2 p` becomes a `float3 p` and the `noise()` function must accept a three-dimensional vector instead. That's all.

Here are some example images of the fractal Brownian motion with different octaves. For the noise function, a Voronoi noise algorithm was used.

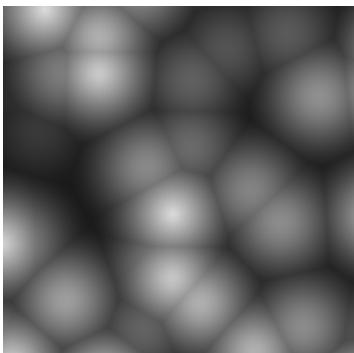


Figure 43: One octave of a 2D Voronoi noise.

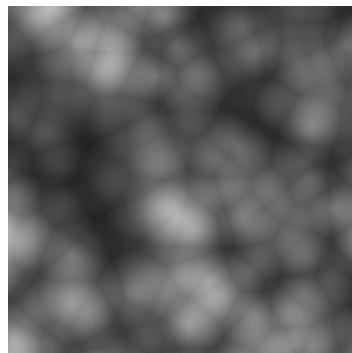


Figure 44: Two octaves of a 2D Voronoi noise.



Figure 45: Three octaves of a 2D Voronoi noise.

It is understandable that with every additional octave, the algorithm has to evaluate the noise at all points again, making it worth considering the impact on performance fractal noise has. However, the final renders look convincingly "cloudy".

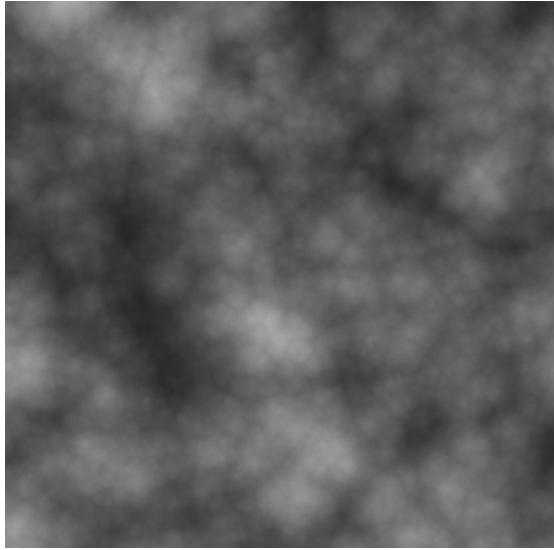


Figure 46: Ten octaves of a 2D Voronoi noise.

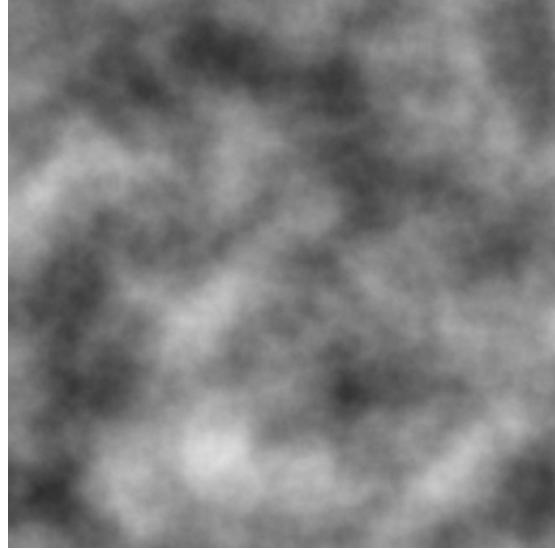
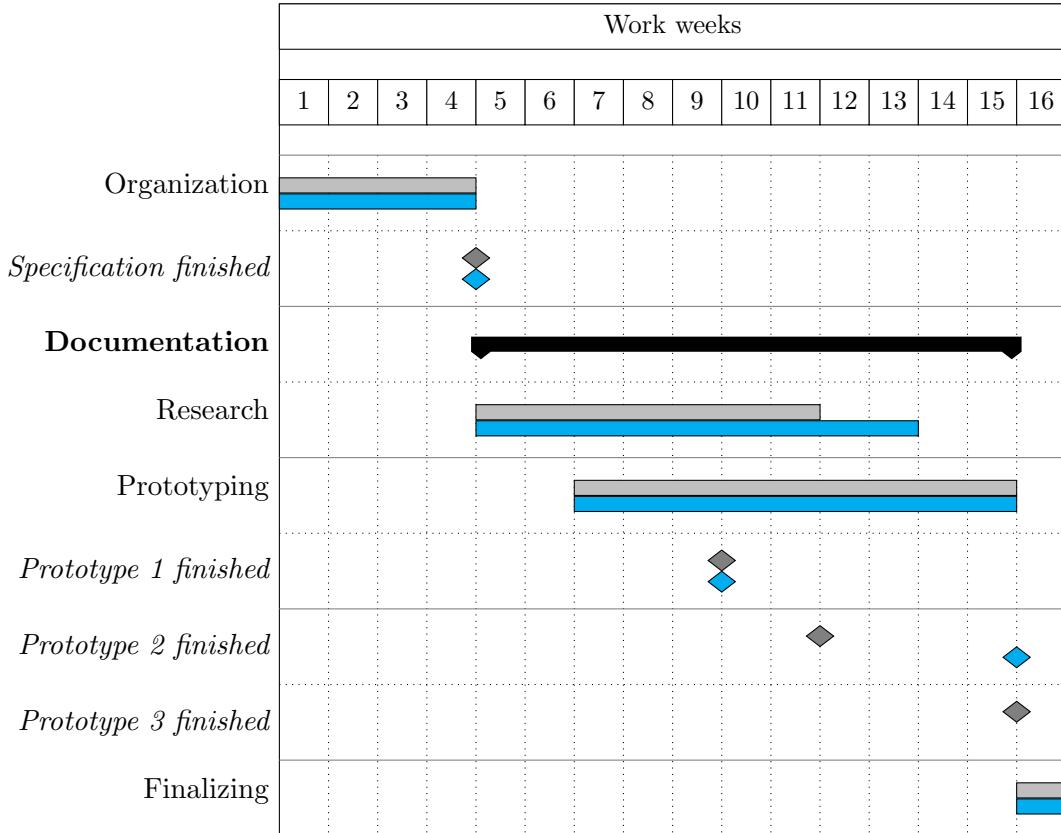


Figure 47: Ten octaves of a 2D Perlin noise.

5 Project Management

5.1 Schedule Adherence

The following chart shows the original schedule (in grey) with a side-by-side comparison with the actual time spent for each task (in blue). It indicates that the schedule was met throughout the project.



As explained in subsection 5.2, "prototype 3" was removed from the schedule, which is why the blue milestone is missing. Due to the unexpectedly freed up development time, the milestone for "prototype 2" was moved to the end of the segment.

It is noteworthy that two more weeks were put into research. This is mainly because new methods and algorithms have been continuously researched during prototyping, which resulted in constant documentation of those findings. It is not regarded as failed planning due to the fact that the extra time gained from the removed prototype allowed for it to happen harmlessly.

Still, the total amount of invested time was about ten percent more than the originally estimated time budget of 128 hours. This is probably due to the fact that there was quite some effort put into the final prototype.

5.2 Goal Discrepancies

Originally, the following three prototypes were planned.

- volumetric rendering
- procedural noise generation
- ray marching

During research and prototyping, it came clear that "ray marching" is in fact a substantial part of volumetric rendering instead of a completely different topic. Hence, only two of the three listed prototypes were implemented. This change led to a significant boost in available development time for the other two prototypes, for which the results could now be fleshed out to a greater extent.

5.3 Future Work

The final prototype leaves a lot to work with for future projects, such as rendering of more cloud genera like the infamous, high-towering cumulonimbus clouds, performance optimizations and much more.

5.3.1 Complete Weather System

Another wonderful idea is to expand the shader into a fully-fledged weather system. Instead of having all those technical parameters, it would instead be dependent on temperature, moistness, altitude, highs and lows, weather fronts and many other meteorological, real variables. It would automatically start to rain when the conditions are met and the cloud shader movement would adjust itself by the wind strength and direction parameters.

5.3.2 Extensive Lighting Features

Despite the considerable effort already put into lighting and illumination methods, there are still some features missing. One of those are *god rays*, the volumetric light shafts that shine through gaps in clouds, giving the scene even more depth. Other absent features are the sun's and moon's halo: A bright circle around the celestial body.

5.3.3 Measure Realism

As described in subsection 6.6, there are several possible approaches to measure the realism of the clouds. This opens up another potential future project.

6 Prototypes and Results

6.1 Preliminary Notes

6.1.1 Completed Prototypes

While researching the topic and experimenting with some dummy shaders, it came clear that "volumetric rendering" and "ray marching" are interchangeable in this matter. Therefore, only two kinds of prototyping have been developed. This change is explained in detail in subsection 5.2.

6.1.2 Dimensions

All of the following documented procedures and algorithms were prototyped and implemented in 3D, but for the matter of explanation, it is described and visualized in 2D.

6.1.3 Unity Variables

The following sections will list code snippets, in which all variables prefixed with an underscore are shader variables exposed to the Unity Editor. This way, they can be changed externally while running the shader code, allowing for convenient debugging. They are from here on out referred to as *parameters*.

6.2 First Draft

The first drafts of prototypes created during this project all revolve around volumetric rendering. Instead of using a signed distance function and evaluating the distance to a 3D data volume, a noise function was used to simulate the *mass* of the cloud. The primary issue was to get the cube transparent where the noise function would return a number close to 0.0 and to color it where the number would be close to 1.0. The approach for solving this issue is done by sampling the cloud's density.

6.2.1 Density Sampling

Like in volumetric rendering, for each pixel fragment, a ray is cast from the fragment into the cube and extended along the view direction for that fragment. Usually, the algorithm can stop for a given ray if the signed distance function returns a small enough distance, meaning the ray has hit a surface of the volume. However, it is different in the case with clouds, where the volume is *translucent* at most points.

To account for that, the ray does not stop until the end of the container cube is reached. It samples the density N times along its path and returns the sum of those samples, giving an approximate qualifier for how dark this fragment should be.

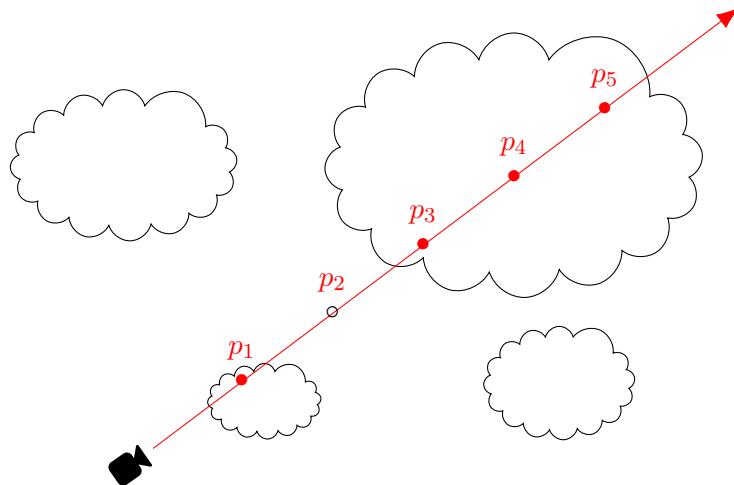


Figure 48: Density sampler ray with $N = 5$.

Understandably, the bigger clouds in Figure 48 represent higher return values of the noise function, meaning denser areas. For the displayed ray, the values for points p_1, p_3, p_4 and p_5 are accumulated and used as a qualifier to color the fragment. In this case, a rather dark tone would be used.

It is notable that N has an exponential impact on the performance, so it should be chosen carefully.

While marching along the ray, the step size is not constant but instead calculated: $d_{step} = \frac{d_{box}}{N}$, where d_{box} is equal to the total distance the ray travels while inside the box. To determine d_{box} , an axis-aligned bounding box (AABB) intersection test [17] has to be done with the container cube and the ray.

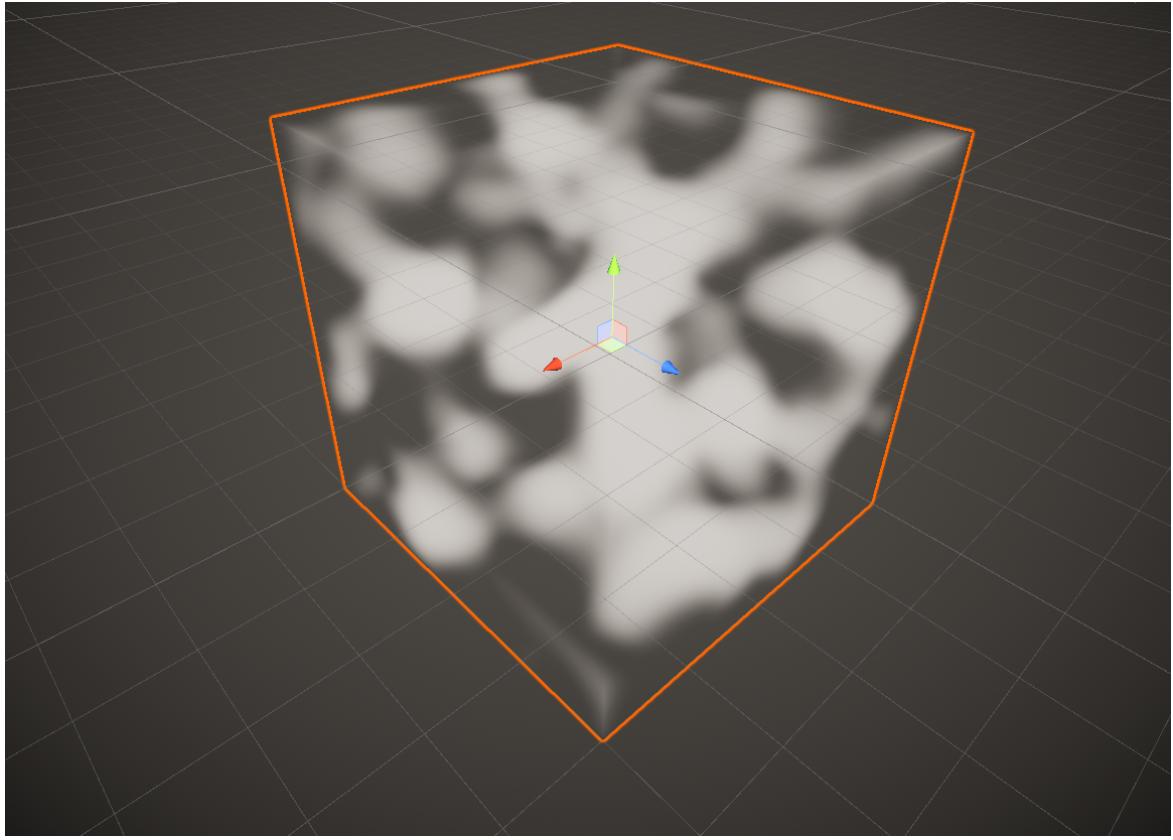


Figure 49: Prototype: Rendered image of sampled density based on 3D Perlin noise.

With this first try, a Perlin noise function was sampled. The returned value had to be normalized in a range of $[0, 1]$ in order for it to be used as alpha value of the color.

6.2.2 Normalizing Density

This is where the exponential function $\exp(x) = e^{-x}$ comes in, which (when clamped between 0 to 1) converts very low values to 1.0 and higher values will converge towards 0.0.

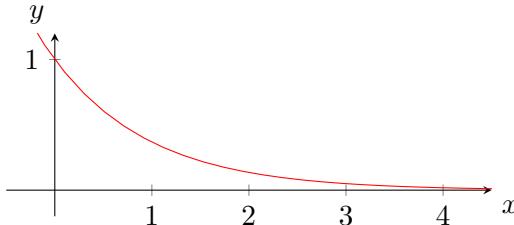


Figure 50: Exponential function $\exp(x) = e^{-x}$.

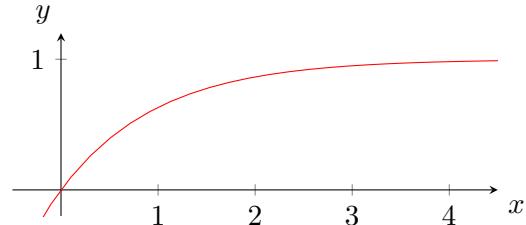


Figure 51: Inverted exponential function $\exp'(x) = 1 - e^{-x}$.

When inverting $\exp(x)$, the function $\exp'(x)$ returns a value that can be directly used for the transparency of the cloud. The denser it gets, the more opaque it will be.

6.3 Improving Noise

After further experimenting with the noise sampling function, the idea arose to combine Perlin and Voronoi noise, which hopefully would create a more distinguished, random pattern. The final sampling function simply multiplies both noise values at a given point `position`, masking them with each other.

```
1 float sampleDensity(float3 position) {
2     float3 vpos = position * _VoronoiScale + _VoronoiOffset;
3     float3 ppos = position * _PerlinScale + _PerlinOffset;
4     float vd = fbmVoronoi(vpos, _VoronoiOctaves, _VoronoiPersistence));
5     float pd = fbmPerlin(ppos, _PerlinOctaves, _PerlinPersistence));
6
7     vd = max(0, vd - _VoronoiDensityThreshold) * _VoronoiDensityMultiplier;
8     pd = max(0, pd - _PerlinDensityThreshold) * _PerlinDensityMultiplier;
9
10    // fixed boost for density by factor 2
11    float density = vd * pd * 2.0;
12    return density;
13 }
```

Listing 15: Implementation of a density sampling function.

By adjusting some of the parameters and increasing the octaves of both noises, a more patchy and cloudy look can be achieved at the cost of performance.

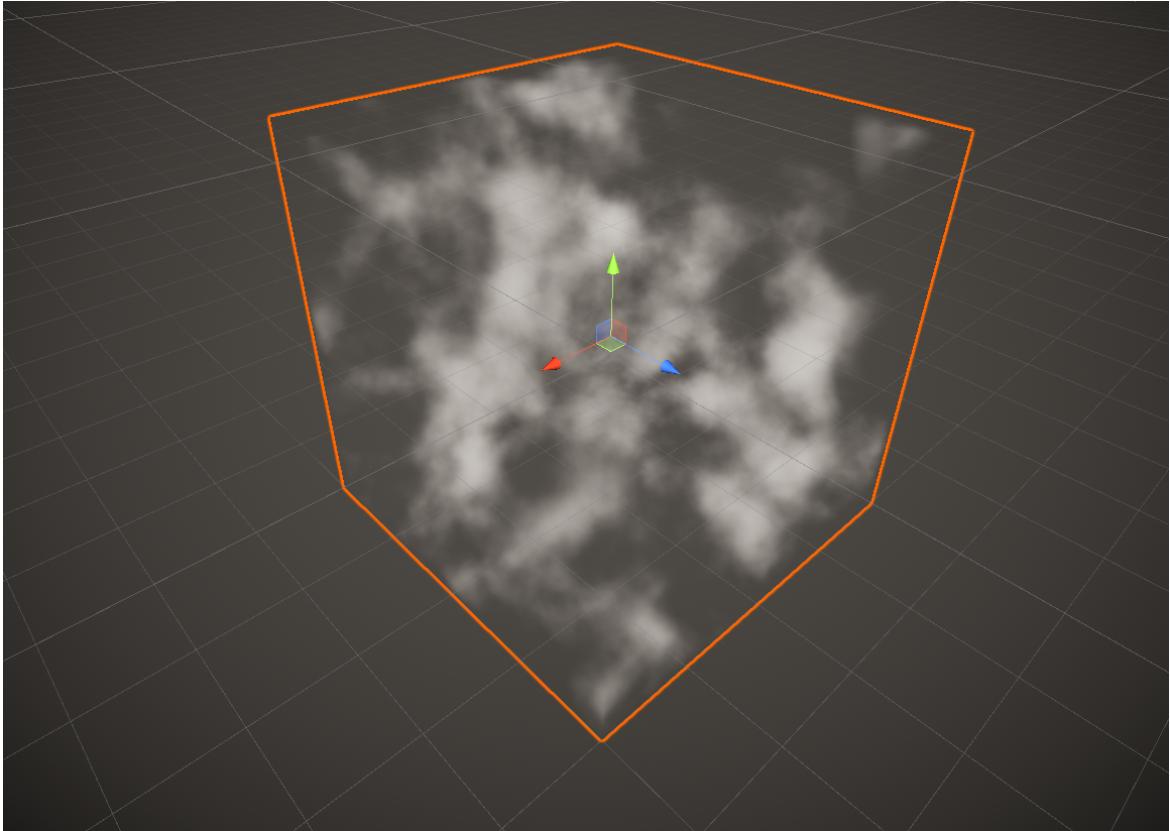


Figure 52: Prototype: Rendered image of sampled density based on mixed noises.

6.4 Light Transmittance and Light Scattering

One of the more prominent lighting features of clouds is its translucency. This phenomenon displays how light bounces and scatters inside the matter, then exits at a different point. This is also called *subsurface scattering* (SSS). It results in illuminated areas where the clouds are thinner, giving it that milky look and "glow" on the outer edge. In nature, subsurface scattering is a very complex and computationally demanding process. For computer graphics however, it is often either simplified or substituted with some other algorithm that produces a similar outcome at lower performance cost.

6.4.1 Sunlight Forward Scattering

When approaching the implementation of subsurface scattering and directional lighting, it seemed most reasonable to start with the sun being visible behind the clouds, or at least shining through them. This implies finding a way to illuminate clouds that cover the sun. In the context of this project, it is called *sunlight transmittance* or *sunlight forward scattering*, since it is not a variant of SSS but rather an approximation.

After some consideration and brainstorming, the following method was chosen to solve the issue:

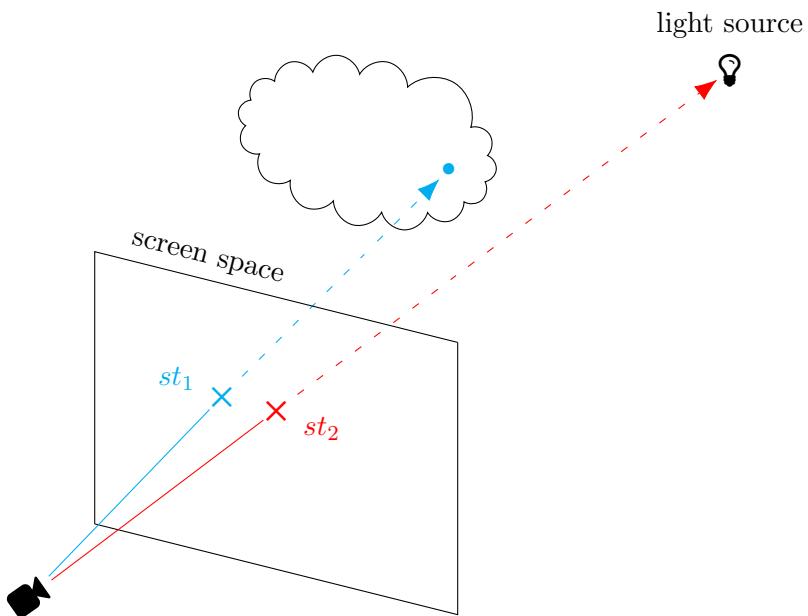


Figure 53: Sunlight transmittance sampling.

When ray casting, both the fragment's and the light source's screen-space position is calculated. Those are two-dimensional coordinates relative to the screen that the camera renders to. Now if the distance $d = \|\overrightarrow{st_1st_2}\| < t$, with t being some threshold, a portion of the sun's color is added to the fragment's color, relative to how small d is.

It is noteworthy that when calculating the screen-space position, the depth value gets lost. Therefore, theoretically, the clouds would be illuminated when $d < t$ even if the sun is in front of the clouds. Given this is almost never the case in games and weather simulations, that particular issue is neglected.

Also, instead of evaluating the distance d , the intermediate angle of both vectors could also be used to avoid calculating the screen-space position, giving $d = \cos^{-1}(\overrightarrow{v_{cloud}} * \overrightarrow{v_{light}})$.

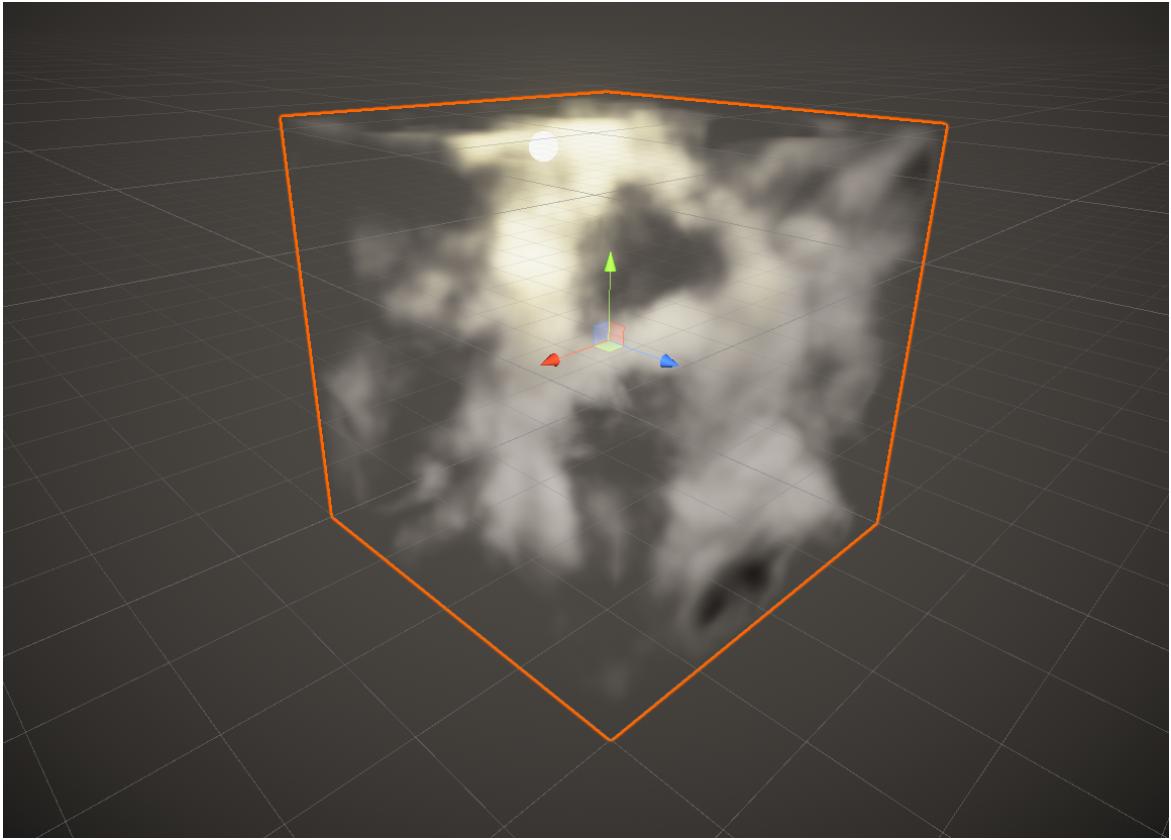


Figure 54: Prototype: Rendered image of sunlight transmittance.

Behind the cube in Figure 54 is a sphere object placed, representing the sun. The sunlight is indeed shining through the clouds, but there are still some minor flaws with the implementation. For example, some clouds are completely illuminated, making them too bright where the cloud would be too dense for the light to pass through.

The following code snippet shows the implementation of the sunlight transmittance mechanism. The `density` variable is the one evaluated in Listing 15.

```

1 float cloudDensity = exp(-density);
2
3 float projectedSunDistance = length(
4     worldToScreenPos(_SunPosition) - worldToScreenPos(worldPosition));
5
6 float sunTransmittance = 1 - pow(
7     smoothstep(0.01, _SunLightScattering, projectedSunDist), _SunLightStrength);
8
9 fixed3 sunColor = sunTransmittance * _LightColor0.xyz * cloudDensity;

```

Listing 16: Implementation of a sunlight transmittance mechanism.

Like in other prototype code listings, there are some parameters to play with. The sunlight strength and its range in screen-space can both be adjusted, for example.

The idea of multiplying by `cloudDensity` on line 9 was to fix the previously described flaw of clouds being too bright.

6.4.2 Directional light

Another challenging part during prototyping was directional light on surfaces facing the sun. Usually in ray marching, a surface normal estimation is done using the gradient. This works well if there is only one point of interest (like a ray-surface intersection point), but as already mentioned before, the ray does not stop sampling points until it reaches the end of the container cube.

So instead of calculating normals for each sample point, another ray is cast from the sample point towards the sun. Along its path, the density is sampled again L times in constant steps. With the lack of an official term, this process is called *lightmarching* in this project.

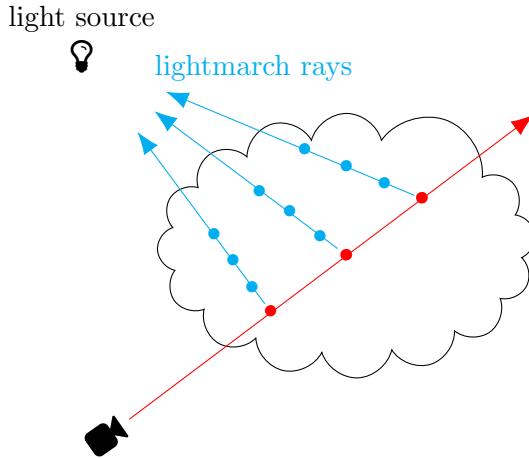


Figure 55: Directional lightmarching samples (part 1).

It is clearly visible that in Figure 55, a lot of density samples return a high value, resulting in a dark fragment color for this ray. To simplify, there is a lot of cloud mass in front of that sample point, so the fragment will not receive a lot of sunlight color.

On the other hand, in Figure 56, only very few samples are even inside a cloud, resulting in a low value overall. This leads to a higher influence of the sun's color for that fragment, meaning the samples are more exposed to the sun.

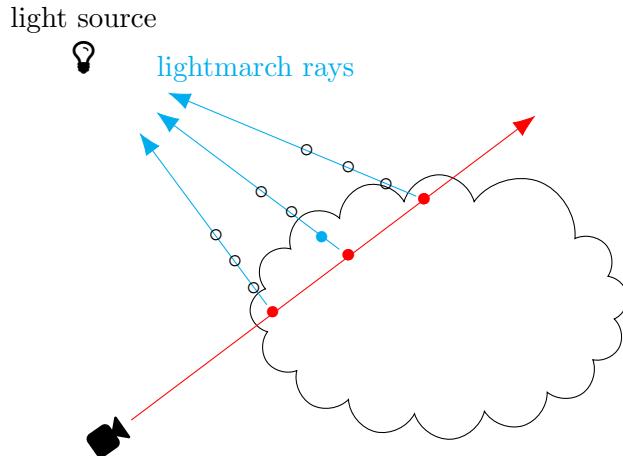


Figure 56: Directional lightmarching samples (part 2).

The implementation for lightmarching is rather straight-forward, given the concept of ray marching is already known.

```

1 float lightmarch(float3 position, float3 direction) {
2     float3 p = position;
3
4     float lightTransmittance = 0;
5     for (int j = 0; j < _MaxLightSteps; j++)
6     {
7         p += direction * _LightStepSize;
8         lightTransmittance += sampleDensity(p);
9     }
10
11    return lightTransmittance;
12 }
```

Listing 17: Implementation of lightmarching.

The method is called during ray marching and the original function is modified like so:

```

1 float2 raymarch(float3 position, float3 direction)
2 {
3     float3 sunDirection = normalize(_SunPosition - position);
4     float lightStepSize = insideBoxDist / _MaxLightSamples;
5     float lightTransmittance = 0;
6
7     [...ray marching...]
8
9     for (int j = 0; j < _MaxLightSamples; j++)
10    {
11        position += direction * lightStepSize;
12        lightTransmittance += lightmarch(position, sunDirection);
13    }
14
15    return float2(density, lightTransmittance);
16 }
```

Listing 18: Implementation of raymarching with lightmarching.

Now, two values are returned instead of just one. Both are later normalized with either $\exp(x)$ or $\exp'(x)$.

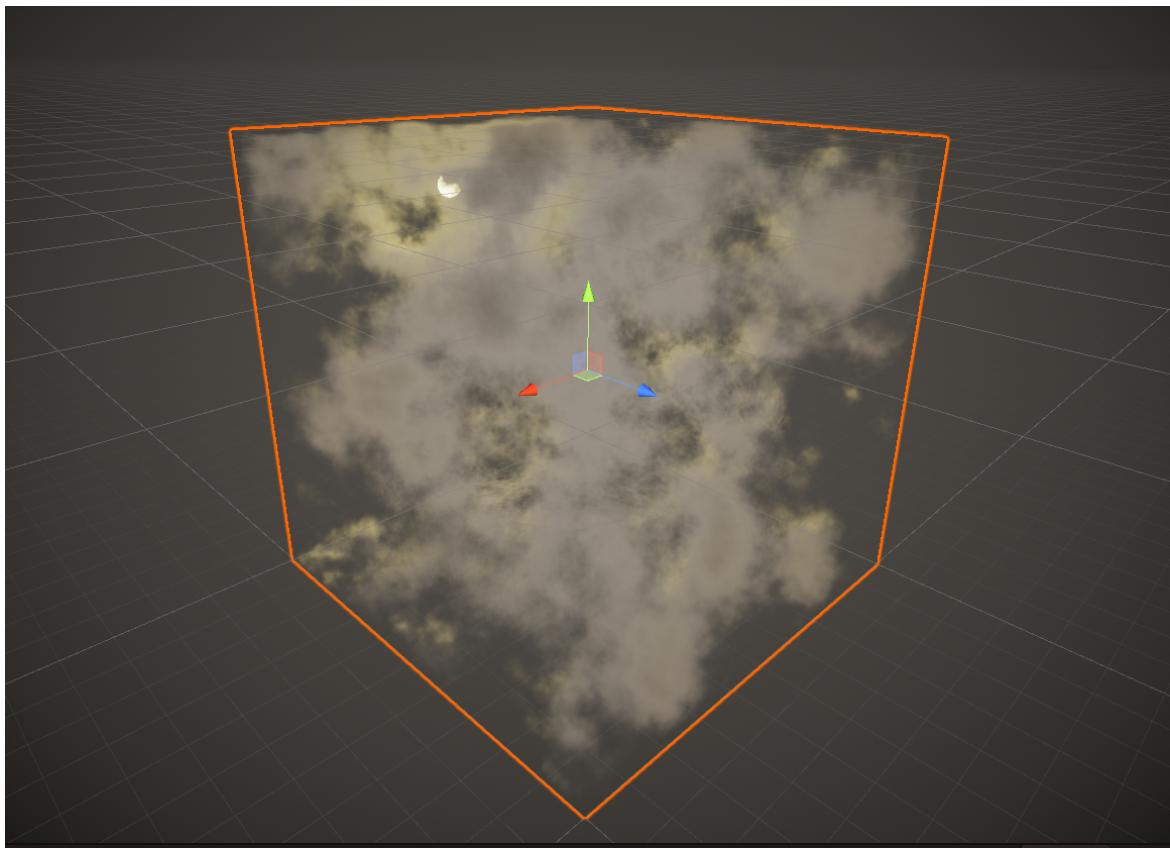


Figure 57: Prototype: Rendered image of directional sunlight implemented with light-marching.

6.5 Final Prototype

All put together and after quite some effort and experimenting, the rendered scene looks quite convincing.

Free assets from the Unity Asset Store were used for trees¹ and rocks².



Figure 58: Prototype: Rendered image of the final prototype (afternoon scene).

¹<https://assetstore.unity.com/packages/3d/vegetation/speedtree/free-speedtrees-package-29170>

²<https://assetstore.unity.com/packages/3d/environments/landscapes/photoscanned-moutainsrocks-pbr-130876>

To demonstrate the capability of the shader in its prototype state, here are some variations of it. There are no code changes in-between the rendered images, the only things that changed are the shader's parameters and Unity Editor lighting settings and colors.

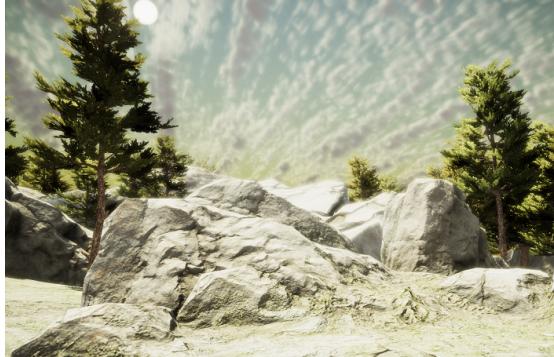


Figure 59: Prototype: Rendered day scene.



Figure 60: Prototype: Rendered puffy sky scene.



Figure 61: Prototype: Rendered night scene.



Figure 62: Prototype: Rendered sunset scene.



Figure 63: Prototype: Rendered clear sky scene.



Figure 64: Prototype: Rendered stormy scene.

6.5.1 Additional Masking

Technically, by multiplying both noise function values in `sampleDensity()`, they already mask each other. In certain types of cloud formations, an additional masking needs to be applied in order to create a cloudscape that does not expand over the whole container cube. This is done by feeding a mask texture into the shader, for which the cube's UV coordinates are used to sample the grey value of the texture at that position.

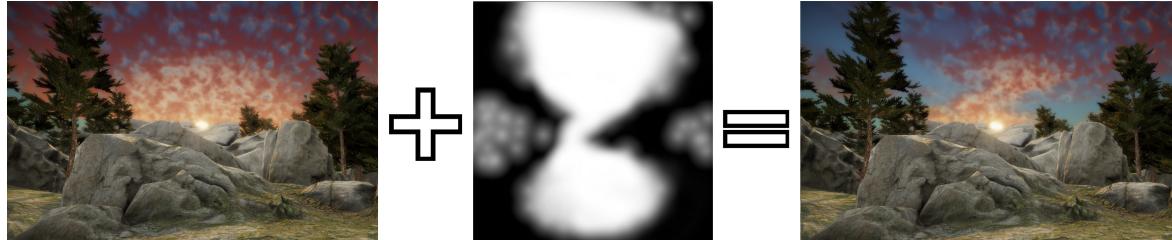


Figure 65: Prototype: Masking with UV coordinates of the container cube.

6.6 Realism Check

While the prototypes in Figure 59 to Figure 64 do look realistically to a certain degree, it is still essential to have some sort of measurable factor with which the rendered images can be compared to real clouds. Some factor that ultimately shows how *real* the rendered clouds actually are, apart from the human eye interpretation.

6.6.1 Objective Interpretation

Before expanding on how to measure the realism of the cloud shader's output, it seems important to objectively identify the capabilities of it first. As originally stated in subsection 2.3, the desired look of the clouds was that of the genus *cumulus*.

In the following subsection, each rendered image is compared to a real-life photograph, putting them into a objectively comparable state to reality.

6.6.2 Real Life Comparison

In all the following comparison images, the left images are photographic references and the right side images are rendered in Unity Engine.

As for the cloud genera, Figure 59 and Figure 60 both resemble cirrocumulus and altocumulus clouds. The cirrocumulus clouds are similar to altocumulus clouds, but they form at higher altitude and are significantly smaller, yet equally puffy. Figure 61 and Figure 62 also show the distinctive cotton-like pattern of the altocumulus genus.

By adjusting the scale of the shader, the clouds can be made smaller. In the following comparison, the directional lightmarching has been turned down to a minimum so the light would look more wan.



Figure 66: Comparison: photographic reference [18] versus the rendered image.

The night-time comparison displays how the different color of the sunlight forward scattering can impact the scene.



Figure 67: Comparison: photographic reference [19] versus the rendered image (at night).

In Figure 62, it is very clearly recognizable that the parameters of the shader can heavily influence the lighting and illumination of the clouds. Unfortunately, the difference of details and density is fairly discernable from reality and the desired appearance is not fully achieved.



Figure 68: Comparison: photographic reference [20] versus the rendered image (at sunset).

Finally, Figure 64 represents clouds of the type *nimbostratus*, which form in low altitude and are dense and dark, often rainy.



Figure 69: Comparison: photographic reference [21] versus the rendered image (stormy).

With the sunlight transmittance being a bright yellow in the rendered image, an attempt was made to make the sun shine through the distant rainy clouds, like in the photographic reference. Since the cloud shader does not end before the horizon, the resulting effect is rather imperceptible.

6.6.3 Convolutional Neural Network

Given there is a convolutional neural network (CNN) that is able to classify images of the sky, the weather or clouds into descriptive labels or even genera of cloud formations, then one could just seed those rendered images into the CNN and verify whether the results are truthfully showing "real" clouds. Of course, this is heavily dependent of how well the CNN was trained.

6.6.4 Generative Adversarial Network

A similar approach to the convolutional neural network is a generative adversarial network (GAN) setup. It describes two neural networks, which contest with each other in a cat-and-mouse game: The *generative* network tries to imitate the training set by generating artificial photographs with many realistic characteristics, while the *discriminative* network tries to tell whether the generated images are fake or not.

With this method, the rendered cloud images could be passed through the discriminative neural network to see if at least the network thinks the images are of real clouds.

6.6.5 Histogram Comparison

The *histogram* is a graphical representation of data like brightness or color distribution of a given photograph. When extracting the color histograms of the real photograph and the one of the rendered image, they could be compared and rated how different in color they are.

6.6.6 Professional Meteorological Assessment

Another viable solution is to let a professional meteorologist inspect and rate the rendered images and judge the realism of the depicted scenarios, which should reveal if the rendered clouds could actually form and exist in reality.

6.7 Performance

With the current implementations of ray marching and lightmarching, the performance of the shader is heavily dependent on the number of samples N and L taken along the rays. With a container cube the size of 400x200 pixels in screen-space (which is a relatively small box of clouds), the GPU load is already enormously large. Here, the number of noise samplings N_{total} is calculated as follows. The values for N and L are both set to a minimum of 25, since it seemed to achieve the best looking results during prototyping.

$$w = 400, h = 200 \\ N = 25, L = 25$$

$$N_{total} = w * h * N * L = 5 * 10^7$$

So for a single frame, the number of times the noise function is called for this example is just about fifty million times. This inconceivably large number makes the current implementation desperately needy of optimization.

6.7.1 Optimization Attempts

6.7.1.1 Compute Shader

The fact that the noise function needs to be sampled so often may not even be the issue, but rather that the noise texture needs to be calculated again each time. Therefore, the idea arose to precalculate the 3D noise volume and feed it to the shader at runtime in the form of a 3D texture cube, hoping that when sampling the same position twice, the noise would only be calculated once.

For this, a compute shader was created with the attempt to generate said texture. Unfortunately, all experiments with compute shaders were unsuccessful, as there is only very little documentation about 3D texture-generating compute shaders for Unity Editor. This attempt was therefore abandoned.

6.7.1.2 Early Exits

For some functions and loops, early exit conditions can speed up the shader. As an example in Listing 18, which describes ray marching combined with lightmarching, the light does not need to be calculated when there is no cloud. The following little extension should fix that issue.

```
1 if (density <= 0) {  
2     return float2(0, 0);  
3 }
```

Instead of checking if `density` is smaller or equal to zero, a very small threshold could be used as well. It is noteworthy that a number larger than zero can create edges on thin clouds, as the abrupt reduction in density may be visible.

Glossary

Ambient occlusion This method darkens points in the scene that are not or only slightly exposed to the light and its environment. 17

Axis-aligned bounding box A non-rotated bounding box enclosing an object completely. 31

Billboard A 2D image always facing towards the main camera. 3

Compute shader A shader which runs on the GPU but outside of the default render pipeline. 46

Constructive solid geometry Short CSG, stands for combining primitive geometric objects with Boolean operators. 16

Convection Convection describes the transfer of heat from movement of liquid or gas. 2

Convolutional neural network A neural network that is able to classify images. 45

Fractal noise In this matter, the same as fractal Brownian motion. 26, 27

Fractal Brownian motion Different iterations of continuously more detailed noise layered on top of each other. i, 26, 47

Generative adversarial network A set of two neural networks, where one generates images and the other tries to tell whether those images are real or generated. 45

GPU Graphics Processing Unit. 46

Gradient The *gradient* denotes the direction of the greatest change of a scalar function. 11

Histogram A graphical representation of data like brightness or color distribution of a given photograph. 45

Lightmarching The same concept as ray marching, but instead of being cast into the volume, it is cast towards the primary light source with a constant step. 36, 37, 38, 43, 46, 52

Low poly A 3D polymesh with a relatively low count of polygons. 4

Noise A randomly generated pattern, referring to procedural pattern generation. i, 20, 29

Noise generation Noise generation is used to generate textures of one or more dimension with seemingly random smooth transitions from black to white (zero to one). 18

Parameters Shader variables exposed to the Unity Editor. 29, 30, 33, 35, 40, 44

Penumbra The partially shaded outer region of diffuse shadows. Also described as soft edges. 14

Polymesh A polymesh is a 3D model composed of polygons or triangles. 4, 11, 51

Procedural Created solely with algorithms and independant of any prerequisites. i, 19, 20, 29, 47

Ray marching Ray marching is a type of method to approximate the surface distance of a volumetric object, where a ray is cast into the volume and stepped forward until the surface is reached. i, 6, 7, 8, 13, 29, 30, 36, 37, 46, 47

Scalar field A scalar field describes a typically three-dimensional grid of elements called *voxels*, each containing a scalar value. 6

Shape blending In SDFs, shapes can be seemingly blended together by returning a interpolated value of those distances. 15

Signed distance function A signed distance function, short SDF, returns a positive distance if the origin is outside the volume and a negative distance if it is inside the volume. 9, 11, 15, 16, 31

Sphere tracing Sphere tracing describes an optimized algorithm of ray marching by using signed distance functions to approximate the surface distance of the volume. 9

Subsurface scattering SSS is a mechanism of light transport in which light enters a translucent object, is scattered around and exits the material at a different point, resulting in illuminated areas where the material is thin. 34

Sunlight forward scattering The process of sunlight shining through and illuminating the clouds which cover the sun. 34, 43, 48

Sunlight transmittance In this matter, the same as sunlight forward scattering. 34, 44

Surface normal A *surface normal* or *normal* is a vector which is perpendicular to a given geometry, like a triangle or polygon. i, 11, 12, 36

Translucent An object or substance that is translucent allows light to be passed through it, meaning it is rendered transparently to some degree. 31

Vector field It is the same as a scalar field, except the voxels are vector values. 6

Volumetric rendering This describes a technique which takes a 3D volume of data and projects it to 2D. It is mostly used for transparent effects stored as a 3D image. i, 6, 11, 29, 30, 31

Voxel Short for *volume element*, a voxel is a value (either a number or a vector) on a scalar or vector field . 6

World space Coordinates defined with respect to a global Cartesian coordinate system. 3

References

- [1] Andrew Schneider, *Real-time volumetric cloudscapes*, <https://books.google.ch/books?hl=en&lr=&id=rA7YCwAAQBAJ&oi=fnd&pg=PA97&dq=real+time+rendering+of+volumetric+clouds&ots=tu16WONk0Z>, [Online; accessed April 2, 2020], 2016.
- [2] Jean-Philippe Grenier, *Volumetric clouds*, <https://area.autodesk.com/blogs/game-dev-blog/volumetric-clouds/>, [Online; accessed April 2, 2020], 2016.
- [3] Jamie Wong, *Ray marching and signed distance functions*, <http://jamie-wong.com/2016/07/15/ray-marching-signed-distance-functions/>, [Online; accessed April 3, 2020], 2016.
- [4] Alan Zucconi, *Volumetric rendering*, <https://www.alanzucconi.com/2016/07/01/volumetric-rendering/>, [Online; accessed April 3, 2020], 2016.
- [5] Patricio Gonzalez Vivo, Jen Lowe, *The book of shaders*, <https://thebookofshaders.com/>, [Online; accessed April 14, 2020], 2015.
- [6] Sebastian Lague, *Lague's youtube channel about shader coding*. [Online]. Available: <https://www.youtube.com/user/Cercopithecus>.
- [7] *Photographic reference of stratus clouds*. [Online]. Available: https://en.wikipedia.org/wiki/Stratus_cloud.
- [8] *Photographic reference of cirrus clouds*. [Online]. Available: https://en.wikipedia.org/wiki/Cirrus_cloud.
- [9] *Photographic reference of an altocumulus cloud formation*. [Online]. Available: https://en.wikipedia.org/wiki/Altocumulus_cloud.
- [10] *Photographic reference of stratocumulus cloudscape*. [Online]. Available: https://en.wikipedia.org/wiki/Stratocumulus_cloud.
- [11] *A polymesh of a cloud*. [Online]. Available: <https://www.utilitydesign.co.uk/magis-metal-mesh-clouds>.
- [12] *Several volumetric cloudscapes from the game horizon: Zero dawn, drawn in real time*. [Online]. Available: <https://tech4gamers.com/horizon-zero-dawn-gets-new-screenshots/>.
- [13] Joey de Vries, *Phong illumination model*, <https://learnopengl.com/Lighting/Basic-Lighting>, [Online; accessed April 8, 2020], 2014.
- [14] Ken Perlin, *Improved noise reference implementation*, <https://mrl.nyu.edu/~perlin/noise/>, [Online; accessed April 27, 2020], 2002.
- [15] ——, *Improving noise*, <https://mrl.nyu.edu/~perlin/paper445.pdf>, [Online; accessed May 04, 2020], 2002.
- [16] Patricio Gonzalez Vivo, Jen Lowe, *Fractal brownian motion*, <https://thebookofshaders.com/13/>, [Online; accessed May 05, 2020], 2015.
- [17] A. Majercik, C. Crassin, P. Shirley, and M. McGuire, *A ray-box intersection algorithm and efficient dynamic voxel rendering*, <http://jcgt.org/published/0007/03/04/>, 2018.
- [18] *Different cloud types*. [Online]. Available: <https://www.gotoknow.org/posts/624816>.
- [19] *Cloudy night sky video*. [Online]. Available: <https://videohive.net/item/clouds-night-sky/15135126>.

- [20] *Altostatus clouds at sunset*. [Online]. Available: <https://www.pinterest.ch/pin/487725834617556668/>.
- [21] *Cloudy and rainy weather report images*. [Online]. Available: <https://www.ekathimerini.com/238896/article/ekathimerini/news/weather-to-turn-cloudy-and-rainy-on-tuesday>.

List of Figures

1	Photographic reference of stratus clouds [7].	2
2	Photographic reference of cirrus clouds [8].	2
3	Photographic reference of an altocumulus cloud formation [9].	2
4	Photographic reference of stratocumulus cloudscape [10].	2
5	The skybox cube as it is used in games.	3
6	The polar sky dome images, folded out.	3
7	A collection of 2D cloud billboards facing the camera.	4
8	The rendered result of the image to the left.	4
9	A polymesh in the shape of an altocumulus cloud [11].	4
10	Several volumetric cloudscapes from the game <i>Horizon: Zero Dawn</i> , drawn in real time [12].	5
11	Ray marching concept visualized.	6
12	Traditional ray marching.	7
13	Traditional ray marching.	8
14	Ray marching with SDF-based sphere tracing.	9
15	Ray marching with SDF-based sphere tracing, without collision.	10
16	Gradient in a 2D scalar field.	11
17	Gradient in a 3D scalar field.	11
18	A 3D cube with a volumetric shader.	12
19	The shaded sphere rendered volumetrically.	12
20	Shadow casting in ray marching.	13
21	Hard shadows only.	14
22	Soft shadows with $k = 7.0$	14
23	Soft shadows with $k = 1.2$	14
24	A blended sphere and box SDF with $k = 0.5$	15
25	A blended sphere and box with holes along each axis.	16
26	Ambient occlusion applied to the scene.	17
27	Only the ambient occlusion part drawn in red.	17
28	Random numbers with the fractional value of sine of x.	18
29	Random numbers with the fractional value of sine of x multiplied by 10000.	18
30	2D random function visualized.	19
31	Perlin grid with pseudo-random gradient vectors.	20
32	Perlin grid with visualized gradient vectors.	20
33	Perlin grid cell with gradient vectors.	21
34	Perlin grid cell with distance vectors from each vertex to the pixel.	21
35	Perlin grid cell with visualized influences of gradient vectors.	21
36	Perlin vertex weights in 2D space with four corners and three interpolations.	22
37	Perlin vertex weights in 3D space with eight corners and seven interpolations.	22
38	2D Perlin noise texture with a 10x10 grid.	23
39	2D Perlin noise texture with Perlin's fade function.	23
40	Voronoi grid with pseudo-randomly assigned seed points for each cell.	24
41	Voronoi grid with seed distances visualized.	24
42	Complete 2D Voronoi noise pattern.	24
43	One octave of a 2D Voronoi noise.	26
44	Two octaves of a 2D Voronoi noise.	26
45	Three octaves of a 2D Voronoi noise.	26
46	Ten octaves of a 2D Voronoi noise.	27

47	Ten octaves of a 2D Perlin noise.	27
48	Density sampler ray with $N = 5$	31
49	Prototype: Rendered image of sampled density based on 3D Perlin noise.	32
50	Exponential function $\exp(x) = e^{-x}$	32
51	Inverted exponential function $\exp'(x) = 1 - e^{-x}$	32
52	Prototype: Rendered image of sampled density based on mixed noises.	33
53	Sunlight transmittance sampling.	34
54	Prototype: Rendered image of sunlight transmittance.	35
55	Directional lightmarching samples (part 1).	36
56	Directional lightmarching samples (part 2).	36
57	Prototype: Rendered image of directional sunlight implemented with light-marching.	38
58	Prototype: Rendered image of the final prototype (afternoon scene).	39
59	Prototype: Rendered day scene.	40
60	Prototype: Rendered puffy sky scene.	40
61	Prototype: Rendered night scene.	40
62	Prototype: Rendered sunset scene.	40
63	Prototype: Rendered clear sky scene.	40
64	Prototype: Rendered stormy scene.	40
65	Prototype: Masking with UV coordinates of the container cube.	41
66	Comparison: photographic reference [18] versus the rendered image.	43
67	Comparison: photographic reference [19] versus the rendered image (at night).	43
68	Comparison: photographic reference [20] versus the rendered image (at sunset).	44
69	Comparison: photographic reference [21] versus the rendered image (stormy).	44

Listings

1	Implementation of a ray march function with constant step.	7
2	Implementation of a volume distance function for a sphere.	7
3	Implementation of a traditional ray march function with converging surface distance approximation.	8
4	Implementation of a signed distance function for a sphere.	9
5	Implementation of ray marching with sphere tracing.	10
6	Implementation of surface normal estimation.	12
7	Implementation of hard shadow casting.	13
8	Implementation of hard shadow casting.	14
9	Implementation of constructive sold geometry.	16
10	Implementation of ambient occlusion.	17
11	Implementation of 2D random number generation.	19
12	Implementation of 3D random number generation.	19
13	Implementation of 2D Voronoi noise algorithm.	25
14	Implementation of fractal Brownian motion function.	26
15	Implementation of a density sampling function.	33
16	Implementation of a sunlight transmittance mechanism.	35
17	Implementation of lightmarching.	37
18	Implementation of raymarching with lightmarching.	37