Intersection Distance Field Collision for GPU

B. Krayer and R. Görge and S. Müller
University Koblenz-Landau, Germany

Abstract
We present a framework for finding collision points between objects represented by signed distance fields. Particles are used to sample the region where intersections can occur. The distance field representation is used to project the particles onto the surface of the intersection of both objects. From there information, such as collision normals and intersection depth can be extracted. This allows for handling various types of objects in a unified way. Due to the particle approach, the algorithm is well suited to the GPU.

CCS Concepts
• Computing methodologies → Collision detection; Mesh geometry models;

1. Introduction
Collision detection is a widely needed and discussed field of research. Various methods exist for different kinds of objects and scenarios, ranging from simulations to gaming. A natural companion to the question of collision detection is distance. A distance field is a function $f_O : \mathbb{R}^n \rightarrow \mathbb{R}$ that returns the closest distance to a given object $O$:

$$f_O(x) = \min_{y \in O} \|x - y\|$$

For a closed object, it is possible to define an inside and outside. Let $\partial O$ be the surface enclosing the object $O$, then the signed distance field (SDF) $s$ can be defined as

$$s_O(x) = \begin{cases} -d_{\partial V}(x), & \text{if } x \in O \\ d_{\partial V}(x), & \text{else} \end{cases}$$

In the following, we will leave out the subscript, if it is clear or irrelevant what object the distance function refers to. Due to the definition 1 the gradient is always normalized.

$$\|\nabla s\| = 1$$

Using this property as the definition of a distance field yields the well-known Eikonal equation. Being able to test for any given point how far away from an object it is and whether it is inside, provides a powerful tool to find collisions between objects.

2. Related Work
Distance fields have a wide variety of representations and applications. Hart [Har96] introduced them in the context of computer graphics with constructive solid geometry operations and the sphere tracing algorithm for efficient rendering.

A common way to store distance fields is a grid of sampled data. While this has advantages, such as ease of use and hardware-accelerated lookup in GPU applications, it requires a large amount of storage to represent high-resolution features.

Frisken et al. [FPRJ00] store the distance field in an octree data structure. Subdivisions of a node are made based on different criteria, such as the reconstruction error at certain positions. Smooth regions can then be stored with less resolution thus saving storage. Bastos and Celes [BC08] extended this data structure to efficiently work on the GPU.

Aside from storing samples, the distance function itself can be approximated with different basis functions. Carr et al. [CBC*01] project an object’s SDF onto radial basis functions, which also closes holes in the data. Koschier et al. [KDB16] use a hierarchical structure, but instead fit polynomials per cell.

Distance fields have been used in many collision detection algorithms. Fuhrmann et al. [FSG03] use distance fields to test particles modeling deformable objects for intersections with another body. Xu and Barbič [XB14] used signed distance fields and sampled points from a mesh to calculate a continuous collision by traversing the path a point takes and checking for intersections. This method was also used in a haptic rendering application [XB17]. Oleynikova et al. [OMT*16] use distance fields as a mapping representation in robotics. This can both be used for the mapping itself and collision-free navigation.

A problem with many pure point sampling approaches is that a simple check if a point lies outside does not give enough information to conclude that the object itself does not intersect with the field. Macklin et al. [MEM*20] solve this problem by solving a constraint optimization problem on the primitives of a shape, such as lines or triangles, to find possible intersections.
3. Distance Fields

The following section will give a brief overview of distance fields in general and the version we use in our algorithm.

3.1. Basics

The SDF has many important properties, as shown in [Har96]. Transformations consisting of a translation $t$, rotation $R$ and a uniform scaling $a$ can be exactly represented as

$$s'(x) = as(\frac{1}{a}R^T(x - t))$$

Using that form, the movement of objects can be incorporated without recomputing the field. From this, it follows, that the gradient transforms simply by inversely transforming the input and rotating the result:

$$\nabla s'(x) = R \nabla s(\frac{1}{a}R^T(x - t))$$

Since $R$ is a rotation matrix, the transformed gradient is still normalized, as stated in equation 3. Constructive solid geometry operations can be very efficiently computed. The operations needed for our algorithm are the intersection $A \cap B$ and union $A \cup B$ of two objects $A$ and $B$.

$$f_{A \cap B}(x) = \max(f_A(x), f_B(x))$$

$$f_{A \cup B}(x) = \min(f_A(x), f_B(x))$$

Since the greatest increase in value occurs orthogonal to the surface, the gradient of the SDF, as with other implicit surfaces, is the normal of the isosurface at the given point. Furthermore, combining the distance and normal information yields a projection operator $\Pi$ that projects a point onto the closest surface. If no unique closest point exists, the gradient and $\Pi$ are not defined.

$$\Pi(x) = x - f(x) \nabla f(x)$$

3.2. Higher Order Fields

While multiple data structures exist to represent distance fields, one of the most well-suited ones for GPU computation is uniform sampling, since the GPU provides hardware-accelerated lookup of interpolated values. If an object feature is smaller than the voxel size of the sampling, all values around it will be positive and thus a zero-crossing will be missed in the interpolation, which would prevent collisions. Increasing resolution may be expensive due to the memory requirements. We chose a similar representation to Bán and Gábor [BV20], but use the Dutch Taylor expansion described in [Kra03, Chapter 4] to correctly increase the order of approximation to second degree. The Dutch Taylor expansion for a distance field is given by:

$$f(x) = x_0 + \frac{1}{2}(x - x_0) \nabla f(x_0)$$

The formula only differs by a factor of $\frac{1}{2}$ from the usual first-order approximation, but when used in a linear interpolation yields a second-order accurate result. As in [BV20] we rearrange the terms to store in a four component texture for 3D points. A voxel element $v_{i,j,k}$ with coordinates $x_{i,j,k}$ is given as:

$$v_{i,j,k} = \left( \frac{1}{2} \nabla f(x_{i,j,k}) \right)$$

$$= \left( \begin{array}{c} a \\ b \end{array} \right)$$

Given an interpolated value $v$ for a point $x$, we can then compute the approximated field value as

$$f(x) \approx x \cdot \pi + \bar{b}$$

$$= \left( \begin{array}{c} x \\ 1 \end{array} \right) \cdot v$$

Section 4.3 will describe the generation and storage of the first-order data and the memory requirements compared to just storing the sampled values. Figure 1 shows a comparison of the zeroth-order and first-order fields.

4. Collision detection

This section will describe the collision detection algorithm and its implementation on the GPU. We don’t require any specific broad-phase detection and thus any algorithm can be used. The only requirement is that a set of potentially colliding pairs of objects is generated.

4.1. Algorithm

For each pair of objects, we determine a region of possible intersection. We choose the intersection of the axis-aligned bounding boxes (AABB) of the two objects, which itself will be an AABB. A possible intersection shape will then be contained in that box. We then place initial collision particles inside of that box.

In many cases, such as two objects just touching at their sides, the intersection box will be very short in one or two directions. To avoid putting too many samples in directions that do not need it, we employ a scheme to distribute the resolutions per dimension according to the length ratios of the intersection box. We start with an initial resolution $N$ for an $n$-dimensional cube, giving $N^n$ particles. Let $b_i, i = 1 \ldots n$ be the $i$-th largest box side length and $r_i, i = 1 \ldots n$ the corresponding resolutions in those dimensions. We compute the ratios $a_i = b_i / r_i$ for $i = 2 \ldots n$. Using the same ratios for the resolutions lets us solve for $r_1$ in the continuous case.

$$\Pi_{i=1}^n r_i = N^n$$

$$r_i \Pi_{i=2}^n a_i = N^n$$

$$r_1 = \frac{N}{\sqrt{\Pi_{i=2}^n a_i}}$$

The other $r_i$ for $i > 1$ can then be computed by using the ratios $a_i$ as $r_i = r_1 * a_i$. For very short dimensions, some of the $r_i$ will result in values less than 1. To counter that, we start from the smallest $r_j$, set it to at least 1 and convert it to an integer by rounding. Afterward, we calculate the resulting number of particles as the product of all $r_i$. We compute the ratio of the actual and desired number $q = \frac{N}{\Pi_{i=1}^n a_i}$. All $r_i, i < j$ will then be multiplied with $\frac{1}{\sqrt{q}}$, which divides the excess resolution equally between the remaining dimensions. This process is repeated for all $j = n \ldots 1$. Algorithm 1 shows this process in 3D.
Instead of testing, whether one of the collision points is inside both objects, we instead project them onto the intersection surface. The general idea of the algorithm is shown in figure 2. According to equations 6 and 8 we define an update step for each collision point $x_i$ as:

$$ x'_i = \begin{cases} 
  \pi_A(x_i) & \text{if } f_A(x_i) > f_B(x_i) \\
  \pi_B(x_i) & \text{else} 
\end{cases} $$

(14)

Due to sampling inaccuracies and discontinuities, we repeat the update step multiple times to arrive at a final position $x'_f$. If an intersection was found $x'_f$ will be on the surface of at least one of the two objects with $f_{A\cap B}(x'_f) < \varepsilon$, where $\varepsilon$ is an iso value to allow for a small offset surface for the collision. $x'_f$ is also possibly inside one of the two objects if they are not just touching. The intersection depth $d_i$ is found by considering the distance to the union of both objects given by equation 7

$$ d_i = f_{A\cup B}(x'_f) $$

(15)

The collision normal $n_i$ from object $B$ to object $A$ is then obtained by the gradient:

$$ n_i = \nabla f_B(x'_f) $$

(16)

By the properties of SDFs, $n_i$ is equal to the normal of the closest point on $B$ and pointing to the outside of the object. Algorithm 2 shows the projection algorithm.

4.2. Accuracy

In contrast to other point-based algorithms, we do not only check for basic containment but search out the collision shape. The projection operator is unique for all points except where no unique closest point exists. The set of such points is a null set and thus nearly all points have a defined projection. An intersection can be found if a sample point lies in the Voronoi region of that intersection. Therefore it is possible to miss features when those regions are not sampled.
This is roughly equivalent to a first-order field with value at that point. This ensures that the normal always points to the voxel center lies on the closest triangle, we return the triangle direction. As the first-order field needs four values instead of the zeroth-order field stored with 32 bit precision at the same resolution. For our textures, thus only requiring twice the memory of a zeroth-order field's resolution.

4.3. GPU

As with other particle-based algorithms, the described procedure is independent for each particle and thus can be efficiently implemented on the GPU. An additional advantage is the hardware-accelerated filtering for textures, which improves the lookup speed of distances and gradients of sampled distance fields. We extend the SDF generation algorithm for meshes from [KM19] to store 

\( v_{ij} \)

according to equation 10. We obtain the gradient as follows.

\[
g \leftarrow \nabla s_0(p)\]

Algorithm 2: Pseudocode for the iterated projection.

Function iterateProj:

Input: Initial point \( p \), SDFs \( s_0, s_1 \), intersection epsilon \( \epsilon \), maximum steps \( maxSteps \)

Output: Projection data

for \( i \) to \( maxSteps \) do

\( d_a \leftarrow s_0(p) \)

\( d_b \leftarrow s_1(p) \)

if \( d_a \geq d_b \) then

\( g \leftarrow \nabla s_0(p) \)

else

\( g \leftarrow \nabla s_1(p) \)

if \( \max(d_a, d_b) < \epsilon \) then

\( \text{result} \leftarrow \{ p, \nabla s_1(p), \max(d_a, d_b), \min(d_a, d_b) \} \)

return true, result

\( p \leftarrow p - \max(d_a, d_b)g \)

Algorithm 3: Pseudocode for the collision detection algorithm.

Input: Array of intersecting pair ids \( P \), Array of accumulated sample particle counts \( C \), Resolutions per particle \( R \), total number of collisions \( n \)

Output: Array of collision points

counts \leftarrow \text{ZeroArray}(n)\)

parallel for \( i \) to \( n \) do

\( idx \leftarrow \text{lowerBound}(C, i) \)

\( a_i, b_i \leftarrow P[idx] \)

bounds \leftarrow \text{getIntersectBounds}(idx) \)

\( p \leftarrow \text{PointInBound}(bounds, idx) \)

\( p, \text{hit} \leftarrow \text{iterateProj}(p, a_i, b_i) \)

if \( \text{hit} \) then

\( \text{atomicInc(counts[idx])} \)

memoryBarrier()

\( \text{totalCount} \leftarrow \text{prefixSum(counts)} \)

return Array \( (\text{totalCount}) \)

parallel for \( i \) to \( n \) do

\( idx \leftarrow \text{lowerBound}(C, i) \)

\( a_i, b_i \leftarrow P[idx] \)

bounds \leftarrow \text{getIntersectBounds}(idx) \)

\( p \leftarrow \text{PointInBound}(bounds, idx) \)

\( p, \text{hit} \leftarrow \text{iterateProj}(p, a_i, b_i) \)

if \( \text{hit} \) then

\( \text{prevEntry} \leftarrow \text{atomicInc(counts[i])} \)

\( \text{result[prevEntry]} \leftarrow \text{hitData} \)

return result

5. Results

We implemented our algorithm in C++ using OpenGL compute shaders for the parallel computations of the intersection finding. Distance fields are accessed via the bindless texture extension. This allows us to directly access textures via an id. A common alternative is using an atlas texture. Times are measured with OpenGL timer queries to accurately get the compute shader execution time. The actual collision handling was delegated to the Bullet Library. For the simulation, the SDF is used to approximate the inertial tensor of

© 2022 The Author(s)

an object. That is accomplished by integrating the formula over all voxels, where a voxel is considered part of the object if its center distance is not positive. Figure 3 shows example scenes computed with our algorithm. The tests were run in Windows 10 with 8GB of RAM, an Intel Core i7-6700HQ CPU, and an NVIDIA GeForce GTX 1070 GPU. We evaluated the algorithm for different scenes. One scene contains a terrain with two models falling on top of each other. The second scene consists of the same terrain, with many objects being dropped on it. The third scene has several balls rolling down a pipe and falling into a bowl. Table 1 shows the results. As all collisions are computed in parallel, we see the average time per colliding object pair going down as the number of those pairs goes up. As the base resolution N corresponds to roughly $N^3$ sample particles, higher values have a large impact on execution time, though the time per pair still stayed in the sub-millisecond range.

6. Limitations
As discussed in section 4.2, intersections can be found, if a point lies in the intersection’s Voronoi region. Thus, depending on the number of sample points and the complexity of the model, small intersections might be missed. A related scenario can be seen in figure 4, where a badly spaced sample does not reach the actual intersection, as the projection has to jump between two nearly parallel lines. This problem is similar to the one encountered in sphere tracing. This might be compensated by using an overrelaxation scheme, such as in [KSK∗14].

7. Conclusion and future work
We presented an algorithm to find the intersection between two arbitrary objects represented by signed distance fields. Because of that, it is not restricted to triangle meshes and does not need any special handling for different types of objects, such as a distinction between convex and concave. Collisions are detected simultaneously for both objects. This is in contrast to previous techniques. These either used a specific geometric representation, such as triangles, or had to check sampling points for each of the two colliding objects separately. Due to its parallel nature, it can be efficiently implemented on the GPU. Many areas could be explored in the future. Currently, collisions are computed independently in each call, but found collisions could be cached for subsequent checks. Instead of uniformly sampling the intersection region, a random or quasi-random sampling might overcome possible aliasing artifacts. To reduce the number of generated collisions, a basic binary grid structure could be used to only allow for one collision in a voxel region. Additionally, a final step could find the k collisions with the deepest penetration and discard all others per object pair.

References


(a) Armadillo model with multiple tori put on around its hands and ears. (b) Multiple different models with different complexity, with interlocking and concave shapes after being dropped on a simple terrain. (c) Multiple balls fall through a winding tube into a bowl.

Figure 3: Examples of different scenes.

<table>
<thead>
<tr>
<th>Scene</th>
<th>$N$</th>
<th>Avg. %collisions</th>
<th>Avg. #pairs</th>
<th>Avg. time [ms]</th>
<th>Avg. time/pair [ms]</th>
<th>Scene image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armadillo and Buddha statue on terrain</td>
<td>5</td>
<td>4.57</td>
<td>3</td>
<td>0.31</td>
<td>0.1017</td>
<td>[Image]</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>9.38</td>
<td>3</td>
<td>0.31</td>
<td>0.1048</td>
<td>[Image]</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>16.18</td>
<td>3</td>
<td>0.38</td>
<td>0.1271</td>
<td>[Image]</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>37.99</td>
<td>3</td>
<td>0.47</td>
<td>0.1554</td>
<td>[Image]</td>
</tr>
<tr>
<td>Terrain with many objects</td>
<td>5</td>
<td>14704.20</td>
<td>1209.40</td>
<td>4.10</td>
<td>0.0034</td>
<td>[Image]</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>118520.59</td>
<td>1325.405</td>
<td>15.27</td>
<td>0.0115</td>
<td>[Image]</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>354805.26</td>
<td>1204.096</td>
<td>37.16</td>
<td>0.0309</td>
<td>[Image]</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>808656.76</td>
<td>1121.437</td>
<td>69.97</td>
<td>0.0624</td>
<td>[Image]</td>
</tr>
<tr>
<td>Pipe with balls and bowl</td>
<td>5</td>
<td>2015.98</td>
<td>408.13</td>
<td>1.72</td>
<td>0.0042</td>
<td>[Image]</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>65590.03</td>
<td>505.23</td>
<td>5.01</td>
<td>0.0099</td>
<td>[Image]</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>288067.48</td>
<td>565.70</td>
<td>17.15</td>
<td>0.0303</td>
<td>[Image]</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>705999.38</td>
<td>576.88</td>
<td>41.21</td>
<td>0.0714</td>
<td>[Image]</td>
</tr>
</tbody>
</table>

Table 1: Timings of the example scenes. All scenarios used a variable time step and were timed using 1000 samples each.

Figure 4: Problematic case for the projection operator. An undersampled region with two objects being close to parallel to each other. Even with a high iteration count, the projection will not converge.