Hardware-Accelerated Collision Detection using Bounded-Error Fixed-Point Arithmetic

Andreas Raabe,[1] Stefan Hochgürtel,[1] Joachim K. Anlauf,[1] Gabriel Zachmann [2]

[1]Technical Computer Science
Bonn University, Germany
{raabe, hochguer, anlauf}@cs.uni-bonn.de

[2]Computer Graphics
Clausthal University, Germany
zach@in.tu-clausthal.de

ABSTRACT

A novel approach for highly space-efficient hardware-accelerated collision detection is presented. This paper focuses on the architecture to traverse bounding volume hierarchies in hardware. It is based on a novel algorithm for testing discretely oriented polytopes (DOPs) for overlap, utilizing only fixed-point (i.e., integer) arithmetic. We derive a bound on the deviation from the mathematically correct result and give formal proof that no false negatives are produced.

Simulation results show that real-time collision detection of complex objects at rates required by force-feedback and physically-based simulations can be obtained. In addition, synthesis results prove the architecture to be highly space efficient. We compare our FPGA-optimized design with a fully parallelized ASIC-targeted architecture and a software implementation.

1 INTRODUCTION

Detecting collisions between a pair of graphical objects is a fundamental task in many areas such as physically-based simulation, automatic path finding, or tolerance checking. Applications are in games, animation systems, and virtual reality, e.g., virtual assembly simulation, or medical training and planning systems.

In most of the applications in these areas, the goal is to avoid collisions, or to enable real-time physically-based simulation. Most approaches today are reactive, i.e., they first place objects at a new trial position, check for collisions, and then compute new forces or positions, based on physical laws, so as to remove any collisions.

This approach demands very efficient collision detection, because it must perform many collision checks per simulation cycle. An emerging application area is the mobile devices market (smart phones, portable games devices). Here, the challenges, besides speed, are size and energy consumption. Another particularly demanding application is force-feedback, where updates of about 1000Hz must be done in order to achieve stable force computations.

Since collision detection is such a fundamental yet challenging task, it is highly desirable to have hardware acceleration available just like 3D graphics accelerators. The benefit is two-fold: a) the system can process objects with higher polygon counts, and b) the system's CPU can be freed from computing collisions.

In this paper, we present a novel, efficient architecture for hierarchical collision detection of two rigid objects. It is based on a novel algorithm for testing a pair of bounding volumes for overlap, which can even be implemented on fixed-point arithmetic. We derive a tight bound on the deviation between overlap testing on fixed-point vs. floating-point arithmetic; this ensures that no false negatives are obtained while producing very few false positives.

We also present an implementation on FPGA hardware along with simulation results concerning its speed and synthesis results concerning its size. Finally, we compare these results with an earlier, parallelized ASIC-targeted architecture, and with a software implementation.

2 RELATED WORK

Considerable work has been done on hierarchical collision detection in software [3, 4, 8, 14, 15]. Some of the bounding volumes (BVs) utilized are spheres, axis-aligned bounding boxes (AABB), oriented bounding boxes (OBB), and discretely oriented polytopes (DOP).

The first work on dedicated hardware for collision detection was presented in [16, 17]. However, they presented only a functional simulation, while we present a RT level implementation along with synthesis results. [12] presented a design that was targeted on ASICs, and was optimized for speed only, and, thus, utilize a total of over 4 million gates and a 756 bits wide bus to a DDR2-RAM. Recently, a commerical hardware was announced that supposedly can do collision detection, among other things [1]. However, no details have been published, in particular, no performance results.

Most other hardware-related research has tried to utilize existing graphics accelerator boards (GPU) [2, 5,

6, 10, 11]. While earlier approaches, such as [11], can basically handle only convex objects, later algorithms, such as [2, 10], have extended these to more general cases such as unions of convex objects or closed objects. The general class of "polygon soups" can be handled by [5], but they use a hybrid approach where the graphics hardware only identifies potentially colliding sets.

All of the approaches using graphics hardware have the disadvantage that they either compete with the rendering process for the same hardware resource, or an additional graphics board must be spent for collision detection. The former slows down the overall frame rate considerably, while the latter would be a tremendous waste, since most of the resources of the hardware would not be utilized at all. Furthermore, most of these approaches work in image space, which reduces precision significantly.

3 BASICS

3.1 Hierarchical Collision Detection with k-DOPs

In this section, we give a short outline of the algorithm of hierarchical collision detection in conjunction with a special kind of bounding volumes, the k-DOP. Additionally, we will quickly recap the separating axis theorem and one of its application, the Separating Axis Test (SAT).

In this paper hierarchical collision detection is utilized to avoid checking every triangle of an object O for intersection with all triangles of object Q. The acceleration data structure is a so-called *bounding volume hierarchy* (BVH), where each leave corresponds to one triangle and inner nodes correspond to groups of triangles. Each node has a bounding volume (BV) attached that bounds all triangles associated with it. In order to achieve a feasible hardware design, we use a binary tree here.

If two objects are checked for overlap, both hierarchies are traversed simultaneously. If their BVs intersect, the next level of BVs is checked. Since two objects will usually intersect only in a very small number of primitives, this yields a significant speed-up in the average case. In practical cases, the complexity is in $O(\log n)$ (n = number of primitives) because only a small diagonal "slice" of constant width down the BVH needs to be visited [9].

In this work, we use k-DOPs as BVs because they were proven to yield very fast collision queries by extensive benchmarking in software [15], and performed very well in our hardware studies [12], too.

k-DOPs are defined over a fixed orientation matrix $\mathbf{D} = (\mathbf{D}_1, \dots, \mathbf{D}_{k/2}, \mathbf{D}_{k/2+1}, \dots, \mathbf{D}_k)$ of vectors in $\mathbb{R}^3$. Each vector $\mathbf{D}_i$ is antiparallel to $\mathbf{D}_{i+k/2}$.

An individual k-DOP is defined by k distances d_i, one along each vector $\mathbf{D}_i$, thus defining a half-space. These DOP coefficients $(d_1, \dots, d_k)$ are the distances of the associated halfspaces to the origin.[1] The $k/2$ pairs of DOP coefficients $(d_i, d_{i+k/2})$ form a so-called *slab* [15].

The intersection of these slabs forms the BV:

$$\text{DOP} = \bigcap_{i=1,\dots,k} H_i, \quad H_i : \mathbf{D}_i \mathbf{x} - d_i \leq 0 \qquad (1)$$

The orientation matrix D, consisting of all the vectors $\mathbf{D}_i$, is fixed and equal for all objects. This yields very memory-efficient description for every k-DOP: only the k coefficients d_i need to be stored.

3.2 Separating Axis Test (SAT)

In this paper we use the so called Separating Axis Test (SAT) introduced by [4, 14].

[4] have shown that two convex polytopes are disjoint if and only if there exists a separating axis orthogonal to a face of either polytope or orthogonal to an edge from each polytope (Separating Axis Theorem).

If only a subset of these axes are tested, false positives might occur, i.e., the polytopes are disjoint while the (incomplete) test yields an intersection. The complete SAT is always correct.

To perform the test, both polytopes must be projected onto each of the candidate separating axes. For each axis, a pair of intervals on that axis results. If one of these pairs is disjoint, then the polytopes must be disjoint (see Fig. 1).

4 EFFICIENT SAT FOR K-DOPS

In this section, we will derive an efficient Separating Axis Test for k-DOPs. Additionally, we will show how the resulting overlap test can be done in fixed-point arithmetic such that no false negatives occur. Finally, we will derive a bound on the deviation of the projection of the fixed-point DOP with respect to the mathematically correct image.

4.1 Precomputation

Since with DOPs the set of vectors $\{D_1, \dots, D_k\}$ is fixed, we can exploit that all possible face orientations of the DOPs within a DOP-tree are the same.

Assume object O is placed relatively to object Q by rotation $\mathbf{M}$ and translation $\mathbf{T}$. Let $\text{DT}(O)$ and $\text{DT}(Q)$ denote the DOP-trees of these objects. As described in Section 3.1, let $(\mathbf{A}_1, \dots, \mathbf{A}_k)$ be the orientations of the DOPs' faces shared by all DOPs in $\text{DT}(O)$ after applying rotation $\mathbf{M}$. Analogously, let $(a_1, \dots, a_k)$ denote the DOP coefficients for DOPs in $\text{DT}(O)$, let $(\mathbf{B}_1, \dots, \mathbf{B}_k)$ denote the vectors shared by all DOPs in $\text{DT}(Q)$, and let $(b_1, \dots, b_k)$ denote the corresponding DOP coefficients.

---

[1] Note that the origin is not necessarily the center of the DOP nor even contained in it.
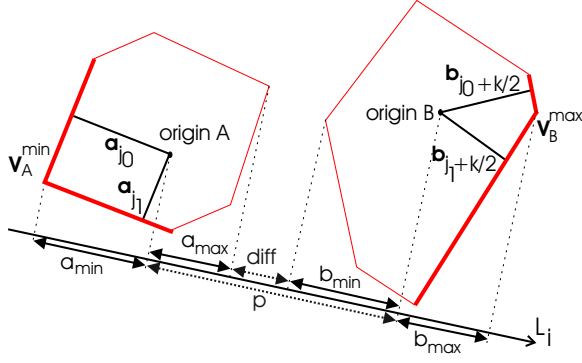
Figure 1: Two DOPs are projected onto test axis L_i. Since their images do not intersect L_i is a separating axis.

Note that everything independent of $(a_1, \ldots, a_k)$ and $(b_1, \ldots, b_k)$ is constant throughout the whole DOP-trees. Hence it can be precalculated at startup to initialize the algorithm (and, later-on, the hardware). Precomputing as much as possible significantly reduces the resulting hardware costs. Since this is done only once per pair of DOP-trees, it is not time-critical.

First, we can precompute the n test axes $\mathbf{L}_i$. All of the following is done for each $\mathbf{L}_i$, so for the sake of simplicity we omit the index i from now on.

Second, the projection $p = \mathbf{L} \cdot \mathbf{T}$ is precomputed.

Third, for each $\mathbf{L}$ a DOP has two vertices $\mathbf{v}_A^{\min}$ and $\mathbf{v}_A^{\max}$ whose projections onto $\mathbf{L}$ have maximum distance. Each of those vertices is formed by the intersection of three faces of the DOP. The correspondences $(j_{A,0}, j_{A,1}, j_{A,2})$ of the orientations whose faces meet in $\mathbf{v}_A^{\min}$ are calculated.

Fourth, and most importantly, in the actual projection

$$a_{\min} = \mathbf{v}_A^{\min} \cdot \mathbf{L}$$
$$= \begin{pmatrix} a_{j_{A,0}} & a_{j_{A,1}} & a_{j_{A,2}} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{A}_{j_{A,0}} & \mathbf{A}_{j_{A,1}} & \mathbf{A}_{j_{A,2}} \end{pmatrix}^{-1} \cdot \mathbf{L}$$

we can precompute the last dot product

$$\mathbf{P}_A := \begin{pmatrix} \mathbf{A}_{j_{A,0}} & \mathbf{A}_{j_{A,1}} & \mathbf{A}_{j_{A,2}} \end{pmatrix}^{-1} \cdot \mathbf{L} \tag{2}$$

$\mathbf{P}_B$ can be precomputed analogously. The mapping vectors for $\mathbf{v}_A^{\max}$ and $\mathbf{v}_B^{\max}$ are $-\mathbf{P}_A$ and $-\mathbf{P}_B$ respectively. This exploits that $k/2$ pairs of DOP orientations are anti-parallel. Note that this is an estimate to the correct solution, since not all possible combinations of DOP-coefficients share all maximum vertices. But it is impossible for any vertex made-up of the intersection of three faces to be inside the DOP, hence only false positives can result.

4.2 Intersection Testing

Using these precomputations, we can project onto the test axes very efficiently:

$$a_{\min} = \begin{pmatrix} \mathbf{a}_{j_{A,0}} & \mathbf{a}_{j_{A,1}} & \mathbf{a}_{j_{A,2}} \end{pmatrix} \cdot \mathbf{P}_A$$
$$a_{\max} = \begin{pmatrix} \mathbf{a}_{j_{A,0}+k/2} & \mathbf{a}_{j_{A,1}+k/2} & \mathbf{a}_{j_{A,2}+k/2} \end{pmatrix} \cdot (-\mathbf{P}_A) \tag{3}$$
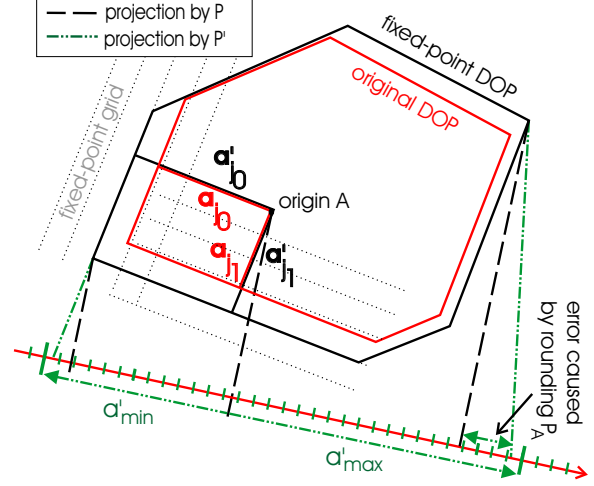

Figure 2: A DOP and its enclosing fixed-point equivalent. Both rounding the DOP to fixed-point numbers and projecting it with $\mathbf{P}'$ instead of $\mathbf{P}$ increases the DOP's image. When checked for intersection false positives can occur.

This is done for $b_{\min}$ and $b_{\max}$ analogously.

The condition for separation is straight-forward now. Let

$$\text{diff}_1 := (a_{\min} + p) - b_{\max}$$
$$\text{diff}_2 := b_{\min} - (a_{\max} + p) \tag{4}$$
$$\text{diff} := \max(\text{diff}_1, \text{diff}_2) \tag{5}$$

then the intervals $[a_{\min}, a_{\max}]$ and $[b_{\min}, b_{\max}]$ are disjoint if and only if $\text{diff} > 0$. And from the Separating Axis Theorem we know that

$$(\text{diff} > 0) \Rightarrow \text{separation}. \tag{6}$$

Eqs. (3)–(6) show the computations that need to be done for each DOP test (and hence cannot be precomputed).

4.3 Fixed-Point Arithmetic

In hardware floating-point arithmetic is very expensive with respect to circuit size and depth. Unfortunately, simply rounding DOP coefficients to fixed-point numbers would result in false negatives, because the intervals on the test axes could become smaller than the projection of the enclosed object. These false negatives are inacceptable, because we might miss collisions. Naïve rounding of the mapping vectors $\mathbf{P}_A$ and $\mathbf{P}_B$ would lead to even more false negatives since distance of the images could be overestimated. Hence we need to round in a manner such that each fixed-point DOP image contains the according floating-point image (see Fig. 2).

First, we need to handle the smaller scale of fixed-point numbers by dividing all DOP coefficients of all DOPs by the largest absolute value of the DOP coefficients in the scenario. This way, 16 bit accuracy still allows for having DOPs the size of a skyscraper and of

a 6mm screw. 36 bit even allow for DOPs the size of the sun and of a screw.

Let rounding of the DOP coefficients to b bits after the point towards $+\infty$ be denoted by $a'_i = \lceil a_i \rceil$. Clearly, the rounded (i.e., fixed-point) DOP contains the original one. Then, $\varepsilon_i = a'_i - a_i$ is the resulting rounding error, with $0 \le \varepsilon_i < 2^{-b}$.

By ensuring that the dihedral angle between all pairs of neighboring faces of a DOP is larger than $\pi/2$, all $P_{A,i}$ are in the interval $[-1, 0]$ [7].[2]

Rounding $\mathbf{P}_{A,i}$ towards $-\infty$ to c bit accuracy results in a rounding error $0 \le \eta_i = \mathbf{P}_{A,i} - \mathbf{P}'_{A,i} < 2^{-c}$.

By simply truncating $\mathbf{P}_{A,i}$, the resulting image would become too small in case of negative DOP coefficients, whereas always rounding up would create the same problem with positive coefficients. Fortunately, we can solve this during calculation simply by adding 2^{-c} to $\lfloor \mathbf{P}_{A,i} \rfloor$ before multiplication with negative DOP coefficients.

Let $\mathbf{a}' := (a'_{j_{A,0}}, a'_{j_{A,1}}, a'_{j_{A,2}})$ and $\mathbf{a}'_k := (a'_{j_{A,0}+k/2}, a'_{j_{A,1}+k/2}, a'_{j_{A,2}+k/2})$. Let $\mathrm{sn}(\mathbf{x})$ be the sum of all $\mathbf{x}_i < 0$.

Then, correct rounding of the images amounts to:

$$a'_{\min} = \mathbf{P}'_A \cdot \mathbf{a}' + 2^{-c}\,\mathrm{sn}(\mathbf{a}')$$
$$a'_{\max} = -(\mathbf{P}'_A \cdot \mathbf{a}'_k + 2^{-c}\,\mathrm{sn}(\mathbf{a}'_k)) \qquad (7)$$

Finally, when computing diff_1, we can simply truncate p to z bits ($p' = \lfloor p \rfloor$). This can create only false positives, because a smaller p' only decreases the apparent distance between the two DOP images. For diff_2 we need to round p up to $\lceil p \rceil$, which, again, can be done efficiently by adding 2^{-z} to $\lfloor p \rfloor$.

Overall, calculating the distances of the fixed-point DOP images amounts to

$$\mathrm{diff}'_1 = (a'_{\min} + p') - \mathbf{b}'_{\max}$$
$$\mathrm{diff}'_2 = \mathbf{b}'_{\min} - (a'_{\max} + (p' + 2^{-z})) \qquad (8)$$
$$\mathrm{diff}' = \max(\mathrm{diff}'_1, \mathrm{diff}'_2) \qquad (9)$$

Now the condition for separation can be given analogously to Eq. 6:

$$((\mathrm{diff}'_1 > 0) \text{ or } (\mathrm{diff}'_2 > 0)) \Rightarrow \text{separation.} \qquad (10)$$

Simulations done early in the design process showed that fixed-point accuracy influences calculation time (Fig. 3). Below 18 bits accuracy, an increasing number of false positives occurs compared to the floating-point implementation and decreases calculation speed. Above 18 bits, a second memory burst is needed to fetch DOP coefficients from DDR-RAM.

4.4 Bound on Fixed-Point Deviation

In this section we will derive a bound on the deviation of the fixed-point image from the mathematically correct image. Let err denote this deviation (called *fixed-point error* in the following)

$$\mathrm{err} := \mathrm{diff} - \mathrm{diff}' \qquad (11)$$

Since diff is defined as $\max(\mathrm{diff}_1, \mathrm{diff}_2)$ (and diff' analogously) we know that

$$\mathrm{err}_1 := \mathrm{diff}_1 - \mathrm{diff}'_1$$
$$\mathrm{err}_2 := \mathrm{diff}_2 - \mathrm{diff}'_2 \qquad (12)$$

$$\min(\mathrm{err}_1, \mathrm{err}_2) \le \mathrm{err} \le \max(\mathrm{err}_1, \mathrm{err}_2) \qquad (13)$$

Inserting Eqs. (3)–(5) and (7)–(9) into Eq. (12) yields

$$\mathrm{err}_1 = (\mathbf{P}_A \cdot \mathbf{a} - \mathbf{P}'_A \cdot \mathbf{a}') - 2^{-c} \cdot \mathrm{sn}(\mathbf{a}')$$
$$+ (\mathbf{P}_B \cdot \mathbf{b}_k - \mathbf{P}'_B \cdot \mathbf{b}'_k) - 2^{-c} \cdot \mathrm{sn}(\mathbf{b}'_k) \qquad (14)$$
$$+ (p - p')$$

and err_2 can be calculated analogously.

To calculate bounds on err_1 and err_2 we need to bound the errors caused by products of mapping vectors and DOP coefficients. Since this is all very similar, we show how it is done, for example, for $\mathbf{P}_A \cdot \mathbf{a} - \mathbf{P}'_A \cdot \mathbf{a}'$.

$$\mathbf{P}_A \cdot \mathbf{a} - \mathbf{P}'_A \cdot \mathbf{a}'$$
$$= (\mathbf{P}_A - \mathbf{P}'_A) \cdot \mathbf{a}' + \mathbf{P}_A \cdot (\mathbf{a} - \mathbf{a}')$$
$$= \sum_{i=0}^{2} (\mathbf{P}_{A,i} - \mathbf{P}'_{A,i}) \cdot \mathbf{a}'_{j_{A,i}} + \sum_{i=0}^{2} \mathbf{P}_{A,i} \cdot (\mathbf{a}_{j_{A,i}} - \mathbf{a}'_{j_{A,i}}) \qquad (15)$$
$$= \sum_{\substack{i=0 \\ \mathbf{a}'_{j_{A,i}} \ge 0}}^{2} (\mathbf{P}_{A,i} - \mathbf{P}'_{A,i}) \cdot \mathbf{a}'_{j_{A,i}} + \sum_{\substack{i=0 \\ \mathbf{a}'_{j_{A,i}} < 0}}^{2} (\mathbf{P}_{A,i} - \mathbf{P}'_{A,i}) \cdot \mathbf{a}'_{j_{A,i}}$$
$$+ \sum_{i=0}^{2} \mathbf{P}_{A,i} \cdot (\mathbf{a}_{j_{A,i}} - \mathbf{a}'_{j_{A,i}})$$

The cross sum of any mapping vector can be interpreted as the image of a vertex of the maximum DOP (all $d_i = 1$). Assume r_{max} to be the greatest distance and $r_{min} = 1$ to be the smallest distance of a vertex of the maximum DOP to the origin. Then $-r_{max}$ is a lower bound and $-r_{min}$ is an upper bound for the cross sum of any mapping vector $\mathbf{P}$.

Let $\mathrm{sp}(\mathbf{x})$ be the sum of all $\mathbf{x}_i \ge 0$ analogously to $\mathrm{sn}(\mathbf{x})$, the sum of all $\mathbf{x}_i < 0$.

With the known boundaries

$$-1 \le \mathbf{a}'_i \le 1 \qquad -2^{-b} \le \mathbf{a}_i\text{-}\mathbf{a}'_i \le 0$$
$$-1 \le \mathbf{P}_{A,i} \le 0 \qquad 0 \le \mathbf{P}_{A,i}\text{-}\mathbf{P}'_{A,i} \le 2^{-c}$$
$$0 \le p\text{-}p' \le 2^{-z}$$

---

[2] This is no hard restriction since every well-constructed DOP should not have acute angles to improve tightness of fit (even for oblong objects in random orientation).

we can bound all summands of Eq. (15):

$$0 \cdot \mathrm{sp}(\mathbf{a}') \leq \sum_{\substack{i=0 \\ a'_{j_{A,i}} \geq 0}}^{2} (\mathbf{P}_{A,i} - \mathbf{P}'_{A,i}) \cdot \mathbf{a}'_{j_{A,i}} \leq 2^{-c} \cdot \mathrm{sp}(\mathbf{a}')$$

$$2^{-c} \cdot \mathrm{sn}(\mathbf{a}') \leq \sum_{\substack{i=0 \\ a'_{j_{A,i}} < 0}}^{2} (\mathbf{P}_{A,i} - \mathbf{P}'_{A,i}) \cdot \mathbf{a}'_{j_{A,i}} \leq 0 \cdot \mathrm{sn}(\mathbf{a}')$$

$$-r_{min} \cdot 0 \leq \sum_{i=0}^{2} \mathbf{P}_{A,i} \cdot (\mathbf{a}_{j_{A,i}} - \mathbf{a}'_{j_{A,i}}) \quad \leq -r_{max} \cdot (-2^{-b})$$

Along with Eq. (15) this amounts to

$$2^{-c} \cdot \mathrm{sn}(\mathbf{a}') \leq \mathbf{P}_A \cdot \mathbf{a} - \mathbf{P}'_A \cdot \mathbf{a}' \leq 2^{-c} \cdot \mathrm{sp}(\mathbf{a}') + r_{max} \cdot 2^{-b} \tag{16}$$

Inserting Eq. (16) into Eq. (14) yields

$$\begin{aligned} \mathrm{err}_1 \leq &\, 2^{-c} \cdot \mathrm{sp}(\mathbf{a}') + r_{max} \cdot 2^{-b} - 2^{-c}\,\mathrm{sn}(\mathbf{a}') \\ &+ 2^{-c} \cdot \mathrm{sp}(\mathbf{b}'_k) + r_{max} \cdot 2^{-b} - 2^{-c}\,\mathrm{sn}(\mathbf{b}'_k) + 2^{-z} \end{aligned} \tag{17}$$

and

$$\begin{aligned} \mathrm{err}_1 \geq &\, 2^{-c} \cdot \mathrm{sn}(\mathbf{a}') - 2^{-c}\,\mathrm{sn}(\mathbf{a}') \\ &+ 2^{-c} \cdot \mathrm{sn}(\mathbf{b}'_k) - 2^{-c}\,\mathrm{sn}(\mathbf{b}'_k) + 0 = 0 \end{aligned} \tag{18}$$

Calculating bounds on err_2 can be done analogously and results in

$$\begin{aligned} \mathrm{err}_2 \leq &\, 2^{-c} \cdot \mathrm{sp}(\mathbf{b}') + r_{max} \cdot 2^{-b} - 2^{-c}\,\mathrm{sn}(\mathbf{b}') \\ &+ 2^{-c} \cdot \mathrm{sp}(\mathbf{a}'_k) + r_{max} \cdot 2^{-b} - 2^{-c}\,\mathrm{sn}(\mathbf{a}'_k) - 0 + 2^{-z} \end{aligned} \tag{19}$$

and

$$\begin{aligned} \mathrm{err}_2 \geq &\, 2^{-c} \cdot \mathrm{sn}(\mathbf{b}') - 2^{-c}\,\mathrm{sn}(\mathbf{b}') \\ &+ 2^{-c} \cdot \mathrm{sn}(\mathbf{a}'_k) - 2^{-c}\,\mathrm{sn}(\mathbf{a}'_k) - 2^{-z} + 2^{-z} = 0 \end{aligned} \tag{20}$$

Since we ensured that the dihedral angles between all pairs of neighboring faces exceed $\pi/2$, r_{max} is bounded by $\sqrt{3}$ [7]. Combining this with Eqs. (17)–(20) and inserting the result in Eq. (14) yields the overall result

$$0 \leq \mathrm{err} \leq \sqrt{3} \cdot 2^{-b+1} + 6 \cdot 2^{-c} + 2^{-z} \tag{21}$$

This gives a bound on the deviation of the image size of a fixed-point DOP with respect to the exact image. Additionally, it formally proves that no false negatives can occur.

5 THE ARCHITECTURE

5.1 The Pipeline

Combining Eqs. (7)–(10) results in the overlap condition

$$\mathbf{P}'_A \cdot \mathbf{a}' + 2^{-c}\,\mathrm{sn}(\mathbf{a}') + \mathbf{P}'_B \cdot \mathbf{b}'_k + 2^{-c}\,\mathrm{sn}(\mathbf{b}'_k) + p' > 0$$
$$\text{or} \tag{22}$$
$$\mathbf{P}'_B \cdot \mathbf{b}' + 2^{-c}\,\mathrm{sn}(\mathbf{b}') + \mathbf{P}'_A \cdot \mathbf{a}'_k + 2^{-c}\,\mathrm{sn}(\mathbf{a}'_k) - (p' + 2^{-z}) > 0$$
$$\Rightarrow \text{separation}$$

Eq. (22) is divided into seven stages to enable pipelining.
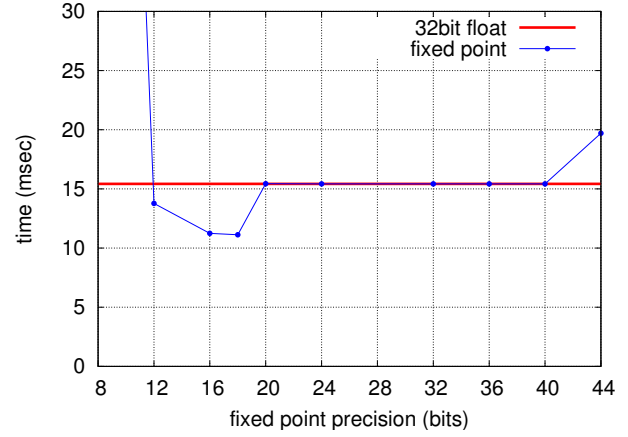

Figure 3: Speed of fixed-point arithmetic for different bit widths. Beyond 18 bits a second, and beyond 40 bits a third memory burst is needed.

Selection. Stage one selects the 12 out of k DOP coefficients defining the outer (maximal) vertices for a given candidate separating axis based on the correspondences (j_A, j_B).[3] Correspondences j_A and j_B contain the indices of 6 of them. The indices of the 6 remaining ones can be derived by simply increasing these indices each by $k/2$. Since we assume wrap around indexing here, this does not need any combinational hardware, but can be done by simply feeding the coefficients into the multiplexers in modified order.

Scalar Products and Fixed-Point Correction. Stages two to five implement the calculation of the scalar products and the fixed-point correction term. So, DOP coefficients have to be multiplied by $\mathbf{P}'$-vector entries and summed up by an adder tree. Additionally, p' ($-(p' + 2^{-z})$ in case of diff'_2) is added. Concurrently, negative DOP coefficients are selected and accumulated. Stage six adds the results of both summations. Multiplying by 2^{-c} is done implicitly by shifting.

Result. Testing $\max(\mathrm{diff}'_1, \mathrm{diff}'_2) > 0$ is done by negating the conjunction of the sign bits.

5.2 Overall Design

The overall architecture is shown in Fig. 4. The calculation is initialized by the host system by sending $(\mathbf{P}'_A, \mathbf{P}'_B, p', j_A, j_B)$ and the addresses of the DOP-trees to the hardware. A controller keeps track of DOP overlap tests that must still be executed and requests the needed DOP coefficients and triangle data. The module "GetData" reads them from memory concurrently to the current calculation. As soon as the parameters are loaded and the last calculation is finished, it feeds them into the pipeline (or the triangle-unit respectively). The pipeline receives not only the DOP coefficients but (from the controller) the data for the next axis test.

---

[3] There are 2 k-DOPs, 2 maximal vertices per DOP, and 3 coefficients defining each vertex (see Section 4.1).
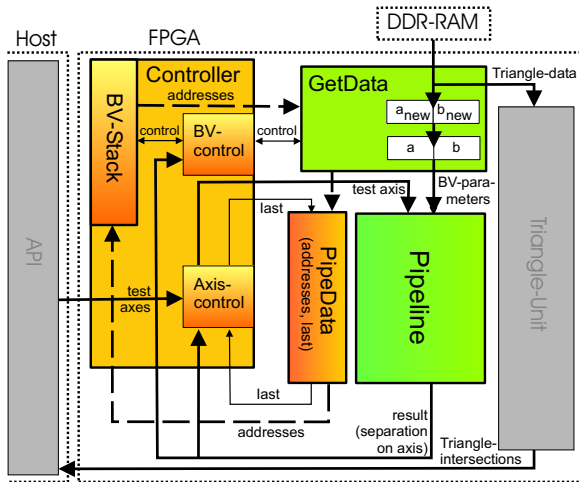
Figure 4: The complete intersection test hardware.

For each DOP pair, n axes are tested. A shift register ("PipeData") holds additional bookkeeping information. For every pipeline stage it contains the indices of the processed DOPs and whether the contained calculation is the last axis test to be executed for the current DOP pair. If this last axis test leaves the pipeline and none of the test axes is a separating axis the controller schedules the child DOPs to be tested. If a separating axis is found, the remaining calculations belonging to the same DOP-pair are obsolete. No new axis tests are initiated and the results of the calculations that are still in the pipeline will be ignored; no new DOP tests are scheduled.

[13] showed that scheduling DOP tests in a stack is far superior to queue control with regards to memory usage. So, as soon as the stack, pipeline, and the Get-Data module are empty, and no intersecting triangles were found, the objects do not intersect and this is reported to the host application. On the other hand, every intersecting pair of triangles is reported to the host immediately.

To check triangles for intersection we utilize the same algorithm that was already proposed in [16] and implemented in VHDL in [13]. It transforms both triangles so that one of them becomes the "unit" triangle. That way, the checks to be performed on the other triangle become very simple and standardized.

5.3 Control

As mentioned in Section 3.2, it is not necessary to test all axes L_i whether they are separating axes. Even more, [14] has shown that it is not efficient to test all axes for OBBs since the probability of the BV to be disjoint decreases rapidly with every non-separating test axis found so far. Fig. 5 shows that this applies for DOPs, too.

On the other hand, we want to eliminate disjoint branches of the DOP trees as early as possible to reduce expensive loading of DOP coefficients. Therefore,
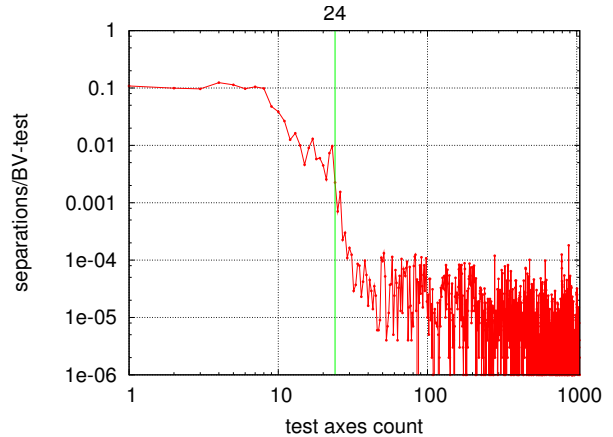

Figure 5: The more axes are tested for intersection the less probable it is for other axes to be separating.
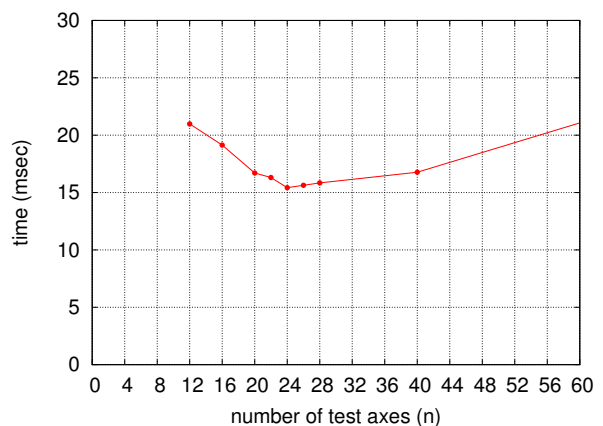

Figure 6: For fixed $k = 24$, our design performs best using $n = 24$ on the target architecture.

we determined which $n \leq N$ gives the best trade-off between axis-testing and parameter-loading. As Fig. 6 indicates, $n = 24$ yields the optimum performance for 24-DOPs and the given memory architecture. 24 axis-tests suffice to test all candidate separating axes generated from the 12 face-orientations of each DOP. Although this exceeds the time to load a complete set of DOP-coefficients (only 20 clock-cycles) by 4 cycles, testing 24 axes seems to reduce the number of false positives enough to yield a performance gain.

Still, there is no reason to stop testing axes if the next DOP-pair is not completely loaded yet. This can happen, for instance, if the memory subsystem is occupied by triangle data. As shown in Fig. 7 continued axes testing until the next set of DOP-coefficients is fetched from memory speeds-up calculation.

6 RESULTS

The target architecture is a Xilinx Virtex II (XC 2V6000, speed grade -4) on an Alpha Data ADM-XRC-II board with 256 MB DDR-RAM at 100MHz connected via a 64 bit wide bus. The FPGA features 144 18-bit multipliers and 6 million gate equivalents. CoCentric from
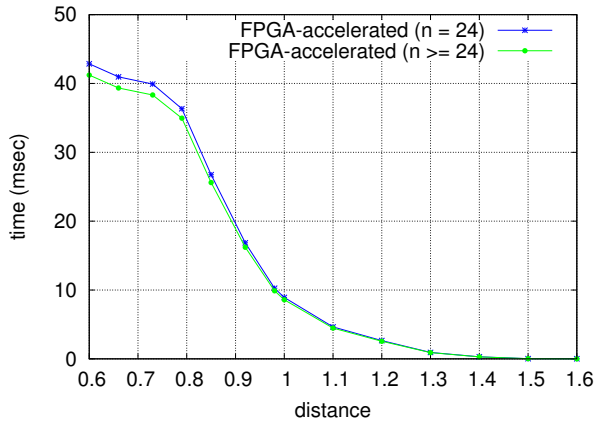
Figure 7: Testing further axes until next DOP-pair is loaded yields a speed-up.

Figure 8: The presented architecture is approximately 4 times faster than a state-of-the-art software intersection test.

Synopsys was used to compile SystemC RTL to VHDL code. Synthesis, Place, Route and Mapping were done with Xilinx ISE 6.3.

6.1 Synthesis Results

Although 19-bit accuracy performs best on our test data with respect to calculation time (Fig. 3), we decided to implement the pipeline for 35 bits fixed-point 24-DOPs to tolerate bigger differences in DOP size (see Section 4.3). Since the target architecture features 18-bit multipliers only, this results in two extra pipeline stages to implement 35-bit pipelined multipliers.

Overall, the pipeline utilizes a total of 7278 out of 33792 slices ($21\% = 1,260,000$ million gate equivalents). Maximum clock frequency is 111.117MHz.

6.2 Benchmarking

All results presented here were obtained with two identical objects (a car headlight) with 5947 triangles [13]. They are placed at different distances from each other and with different rotations. For each constellation, the time to detect all intersecting triangles is determined. Fig. 8 shows the comparison of our new architecture with a state-of-the-art software intersection test running on a 1 GHz Pentium III with 512 MByte main memory. Memory bandwidth and speed are identical on both systems and hence allow for a direct performance comparison. The presented acceleration hardware yields a speed up of about factor 4.

7 CONCLUSION AND FUTURE WORK

We have presented a novel algorithm for hierarchical collision detection of pairs of virtual objects. We have also presented a highly space-efficient, FPGA-optimized architecture implementing this algorithm on an FPGA using fixed-point arithmetic. The fixed-point calculations do not produce any false negatives, and we have given bounds on the deviations from floating-point arithmetic.
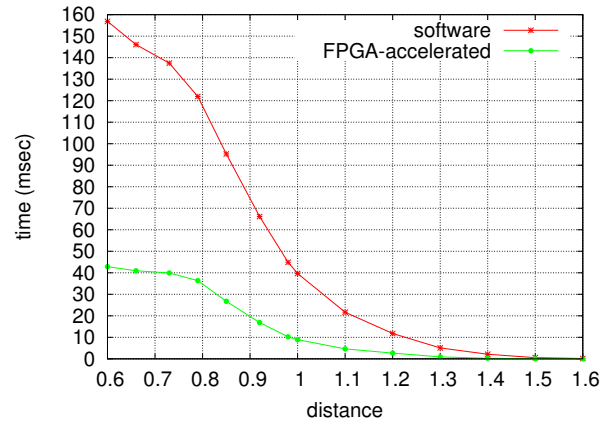
Simulation results for collision queries using this architecture proved that a speed-up of 4 compared to state-of-the-art software intersection tests on a standard CPU can be obtained. Taking earlier ASIC-targeted results into account [13], we conclude that an ASIC implementation of our novel algorithm and architecture will perform by one or two orders of magnitude faster than a software implementation and even the FPGA-implementation. Synthesis results proved the design to be highly space-efficient.

In addition, our novel DOP overlap test algorithm lends itself well to parallelization. Only a slight modification of the controller is necessary to use multiple pipelines to test multiple candidate separating axes in parallel. In conjunction with its low area consumption, this allows for an easy implementation of a highly parallelized architecture with an expected speed-up linear to the number of pipelines.

Here, memory bandwidth becomes the limiting factor for speed of collision queries. Possible solutions could be compression of DOP coefficients and the introduction of a cache.

Another important topic is fixed-point accuracy. Here, a lot of different ways to get smaller projections are conceivable.

Collision detection of deformable objects is another important issue. It remains an open problem, which algorithms and data structures are best suited for hardware implementation. Furthermore, we will evaluate different kinds of primitives like quadrangles and NURBS.

REFERENCES

[1] Ageia. White paper, May 2005. `http://www.ageia.com/pdf/wp_2005_3_physics_gameplay.pdf`.

[2] George Baciu, Wingo Sai-Keung Wong, and Han-qiu Sun. RECODE: an image-based collision de-

tection algorithm. *The Journal of Visualization and Computer Animation*, 10(4):181–192, October - December 1999. ISSN 1049-8907.

[3] Jens Eckstein and Elmar Schömer. Dynamic Collision Detection in Virtual Reality Applications. In *Proc. The 7-th Int'l Conf. in Central Europe on Comp. Graphics, Vis. and Interactive Digital Media '99 (WSCG'99)*, pages 71–78, Plzen, Czech Republic, February 1999. University of West Bohemia.

[4] Stefan Gottschalk, Ming Lin, and Dinesh Manocha. OBB-Tree: A Hierarchical Structure for Rapid Interference Detection. In *SIGGRAPH 96 Conference Proceedings*, Holly Rushmeier, Ed., pages 171–180. ACM SIGGRAPH, Addison Wesley, August 1996. held in New Orleans, Louisiana, 04-09 August 1996.

[5] Naga K. Govindaraju, Stephane Redon, Ming C. Lin, and Dinesh Manocha. CULLIDE: Interactive Collision Detection Between Complex Models in Large Environments using Graphics Hardware. In *Graphics Hardware 2003*, pages 25–32, July 2003.

[6] Alexander Gress and Gabriel Zachmann. Object-Space Interference Detection on Programmable Graphics Hardware. In *SIAM Conf. on Geometric Design and Computing*, M. L. Lucian and M. Neamtu, Eds., pages 311–328, Seattle, Washington, November13–17 2003. Nashboro Press.

[7] Stefan Hochgürtel, Andreas Raabe, Gabriel Zachmann, and Joachim K. Anlauf. Collision Detection for k-DOPs using SAT with Error Bounded Fixed-Point Arithmetic. Tech. rep., University of Bonn, September 2005. `http://www.collisionchip.de`.

[8] P. M. Hubbard. Collision detection for interactive graphics applications. *IEEE Transactions on Visualization and Computer Graphics*, 1(3):218–230, September 1995. ISSN 1077-2626.

[9] Jan Klein and Gabriel Zachmann. The Expected Running Time of Hierarchical Collision Detection. In *SIGGRAPH 2005, Poster*, Los Angeles, August 2005. `http://www.gabrielzachmann.org/`.

[10] Dave Knott and Dinesh K. Pai. CInDeR: Collision and Interference Detection in Real-Time Using Graphics Hardware. In *Proc. of Graphics Interface*, Halifax, Nova Scotia,Canada, June11–13 2003.

[11] Karol Myszkowski, Oleg G. Okunev, and Tosiyasu L. Kunii. Fast collision detection between complex solids using rasterizing graphics hardware. *The Visual Computer*, 11(9):497–512, 1995. ISSN 0178-2789.

[12] Andreas Raabe, Blazej Bartyzel, Joachim K. Anlauf, and Gabriel Zachmann. Hardware Accelerated Collision Detection — An Architecture and Simulation Results. In *Design Automation and Test (DATE)*, Munich, Germany, March7–11 2005. `http://www.gabrielzachmann.org/`.

[13] Andreas Raabe, Blazej Bartyzel, Joachim K. Anlauf, and Gabriel Zachmann. Hardware Accelerated Collision Detection — An Architecture and Simulation Results. In *Design Automation and Test (DATE)*, Munich, Germany, March7–11 2005. `http://www.collisionchip.de`.

[14] Gino Johannes Apolonia van den Bergen. *Collision Detection in Interactive 3D Computer Animation*. PhD dissertation, Eindhoven University of Technology, 1999.

[15] Gabriel Zachmann. Rapid Collision Detection by Dynamically Aligned DOP-Trees. In *Proc. of IEEE Virtual Reality Annual International Symposium; VRAIS '98*, pages 90–97, Atlanta, Georgia, March 1998.

[16] Gabriel Zachmann and Günter Knittel. An Architecture for Hierarchical Collision Detection. In *Journal of WSCG '2003*, pages 149–156, University of West Bohemia, Plzen, Czech Republic, February3–7 2003. `http://www.gabrielzachmann.org/`.

[17] Gabriel Zachmann and Günter Knittel. High-Performance Collision Detection Hardware. Tech. Rep. CG-2003-3, University Bonn, Informatikk II, Bonn, Germany, August 2003. `http://www.gabrielzachmann.org/`.