Broad Phase Collision Detection: New Methodology and Solution for Standardized Analysis of Algorithms

Ygor Rebouças Serpa¹, Maria Andréia Formico Rodrigues (Orientadora)
Programa de Pós-Graduação em Informática Aplicada (PPGIA)
Universidade de Fortaleza (UNIFOR) – 60811-905 – Fortaleza-CE – Brasil
{ygor.reboucas,andreia.formico}@gmail.com

Abstract—Collision detection is a computational problem focused on the identification of geometric intersections between objects and, in general, proximity relationships among them. Despite its notorious relevance and applications in various computing fields, few authors have proposed solutions that are both general and scalable. Additionally, until the time of publication of the results of this work, there was no standard methodology for the analysis of algorithms, neither in academia nor in the industry: only proprietary scenes and comparative studies had been developed, making it difficult to reproduce and compare results. To tackle the issues previously mentioned, we present a new general and scalable solution for broad phase collision detection and a new methodology for comparative analysis of algorithms, named Broadmark, whose open-source code is publicly available, with the goal of transferring knowledge to academia, industry, and society, so far lacking in the scientific literature. Thus, by doing so, we aim to contribute to the generation of robust and multi-faceted solutions applied to various scenarios and, consequently, to greater transparency, ease of modification/extension and reproducibility of results.

I. INTRODUCTION

Collision detection can be seen as a generalization of the k-Nearest Neighbour problem for dynamic scenes with non-punctiform objects, introducing complexities such as the shape and behavior of objects [1]. Although collision detection has been studied for many years, the field still features relevant and active research that has not yet been resolved, introducing additional challenging problems, such as simulation of massive dynamic scenes, deformable bodies, solid-liquid coupling, and robust constraint solving, among others [2], [3]. As a whole, despite its large success, the task is known to be unstable at massive scales and to vary widely in performance whenever scenes deviate from their expected states [4].

Over the years, several efficient solutions have been proposed [5]–[8], however, most works employ a limited set of scenes and algorithms for their comparisons, failing to provide a strong foundation to support their claims. Collectively, these works lack a shared representative methodology, an effort attempted only by Woulfe and Manzke [9], to limited success. On top of that, many authors sacrifice generality over scalability, narrowing their solution to either the static case, in which only a fraction of objects moves, or the dynamic one, in which all of them do. Combined, these problems reveal how challenging it is to weigh the strengths of each work and to faithfully reproduce their results, concerns of paramount importance given the current reproducibility crisis [10].

This work addresses the main research questions fully detailed in the M.Sc dissertation [11] available at (https://1drv.ms/b/s!Aq35PBOZWmsjppQxK3utjL4qDvLfA):

1) To propose an open and extensible standard methodology, yet non-existing, to the development and study of broad phase collision detection algorithms [12],
2) To create a both general-purpose and scalable novel algorithmic solution to the broad phase collision detection field [4], and
3) To make all source code of the developed tools necessary for this research and evaluation of algorithms, publicly available on the GitHub repository (https://github.com/ppgia-unifor/Broadmark), so that anyone interested can inspect, learn from it, test it, develop it, and build on it, thus, in an effort also contributing to the transfer of the knowledge base to the academy, industry and society.

More specifically, we have developed the Broadmark methodology [12], a research development environment containing 12 algorithm implementations, plus variants, for both CPU and GPU (Table I), as well as three standardized testing scenarios (Figure 1), representatives of the static, dynamic, and uniform cases, with either same sized or randomly sized objects [12]. As part of this system, we have also developed a novel hybrid and adaptive solution based on KD-trees, the Sweep-and-Prune (SAP) algorithm, and the incremental detection paradigm, competitive on all tested scenarios [4].

II. RELATED WORK

To date, the dominating approach for assessing collision detection algorithms is to design your own environment. This includes defining the test cases, searching for baseline algorithms and timing each solution [5]–[8], [13]–[21]. This methodology has a number of flaws, to name a few, results are difficult to reproduce, baselines are not state-of-the-art, and no external guarantees of fairness are given. For instance, the solutions proposed by Lo et al. [13], Avril et al. [17] and Liu et al. [6], although similar and contemporary, share almost nothing in their validation strategies. This problem is relevant even within a same institution [17]–[20]. Regarding fairness, it is not uncommon for results to be presented under favorable situations, such as grid-based algorithms on homogeneously distributed scenes [6], [13] or temporal-optimized algorithms on nearly static scenes [5], [6], [16]. Conversely, algorithms devoid of temporal optimizations are usually tested on fully dynamic scenes only [13], [21], [22]. The work of Woulfe and Manzke [9], presented in 2009, is the most recent attempt to introduce an open space for collision detection research.

¹M.Sc. dissertation.
However, it did not get traction within the community and, as of writing, is unfortunately no longer available.

In this context, the Broadmark system [12], presented in the M.Sc. dissertation, is meant to (1) reduce the barrier-of-entry to the field by exposing a set of ready-to-use tools, (2) be a repository of state-of-the-art collision detection algorithms, and (3) benchmark a wide range of solutions on a representative set of scenes. As of writing, the system includes algorithms from five different sources that span single and multi-threaded algorithms on both CPU and GPU, shown in Table I. Additionally, representative scenes of the most frequently seen setups in the literature are also bundled: the nearly static (Figure 1.a), the uniformly distributed (Figure 1.b) and the fully dynamic (Figure 1.c), up to a million objects. Finally, the system is designed towards extensibility, featuring a common interface for all new algorithms and a Unity [23] based editor for authoring new massive scenes.

The Broadmark framework, in its early stages, revealed that most solutions are biased towards one of the three developed scenarios, with none being significantly competitive on all three simultaneously [4]. For a solution to be efficient on all these three scenes, we hypothesized that it had to (1) have temporal optimizations, for the static case; (2) not be impaired by these optimizations on dynamic scenes; and (3) be flexible enough to handle uniformly and non-uniformly distributed objects similarly.

To jointly address these three propositions, we have developed an adaptive hybrid algorithm based on an efficient KD-tree with a SIMD-optimized SAP implementation. Both algorithms run under the incremental detection regime, for the temporal optimization, but can be adaptively toggled to perform the full detection when a specified threshold is met. While KD-trees, SAP and incremental detection, in isolation, are well-established techniques, the synergy between them had not yet been explored in the literature. In the series of tests we carried out, our solution out-performed well-known state-of-the-art algorithms [6], [7], [14], [22] in all three scenarios, being competitive among multi-core and GPUs in several instances.

III. BROADMARK

The Broadmark system is composed of two independent modules: (1) the simulation generator, developed using the Unity game engine, and (2) the algorithm runner, developed purely using the C++ language. The former is accessible either via a pre-compiled wizard, designed for those with no Unity expertise, or via the Unity project itself, through the engine’s editor, while the latter is a command-line tool, for maximum flexibility. As an extra convenience, we also provide a Python tool to design large benchmark schedules.

A. The Simulation Generation

The simulation generation tool was designed using Unity to ease the creation and maintenance of massive scenes with beautiful real-time visualizations. When ran, simulations are baked to disk in binary format, completely decoupling the scene generation from the benchmark. Thus, the time spent on generating scenes is constant with regard to the number of algorithms. Moreover, it ensures that each test uses the exact same Axis-Aligned Bounding Boxes (AABBs), improving fairness. Finally, we have designed the system to modulate each physical aspect of the scene to the number of objects used. This way, the ratio between the scene volume and the total volume occupied by objects is constant, regardless of how many objects are used.

Within this system, three scenes have been developed: Free Fall, Brownian, and Gravity, representatives of the static, uniform and dynamic cases, respectively, and designed to be run from one thousand objects to a million. Figure 1 shows the three scenarios with four thousand same sized objects (left half, in green) and varied-sized objects (right half, in assorted-colors).

B. The Algorithm Runner

In the second module, we have included everything related to running and analyzing the algorithms, as well as the scene reader and the logging functionality. Within this module, we gathered 12 sets of algorithms, including original implementations and known algorithms from the literature and industry, spanning serial, parallel, CPU, and GPU algorithms [4], [6], [7], [14], [22]. Table I enumerates each set of algorithms and their most important features. While some sets have only one algorithm, others, such as BF and SAP, include several variants, for instance, SIMD, parallel, and SIMD + parallel implementations. Other families, such as the DBVT family, have unique features, such as being able to run on a forward pass (DBVT F) or in deferred mode (DBVT D), respectively, optimized for the static and dynamic cases.

IV. HYBRID ALGORITHM

To jointly satisfy the criteria of generality and efficiency, we hypothesized that three properties are needed: flexibility, to be versatile enough to handle different object distributions; efficiency, to be able to prune huge portions of the search space quickly; and adaptivity, to avoid worst-case scenarios and bias towards one kind of simulation over another. In practical terms, to be simultaneously efficient on both static and dynamic scenes.

For the first and second concerns, we have developed a two-tier approach based on the KD-trees and the SAP technique. The former is responsible for dividing the set of objects into smaller, independent, groups, and the latter for processing each individual subproblem using sorting, yielding the set of colliding objects for each sub-problem. Both algorithms complement each other, as the KD-tree is allowed to be shallower, mitigating its overhead, and the SAP is performed over smaller sets, reducing its \(O(n^2)\) time complexity. To efficiently build and update this structure, we have developed an update algorithm which runs in linearithmetic time, is idempotent, and is able to work efficiently for both minor adjustments and major tree changes, thus,
being optimized for both static and dynamic scenes. Orthogonal to these developments, we explored the use of SIMD instructions to accelerate the SAP algorithm and an efficient memory layout custom-tailored to reduce the costs of the most expensive operations carried out during the tree update algorithm.

To efficiently cope with static scenes, it is paramount to have algorithms whose time complexity is dependant on the number of dynamic objects, instead of the total. To introduce reasoning within our algorithm, we employ the incremental detection, which consists in detecting only new collisions and ceased collisions. For this to work, we start from the set of collisions from the previous frame, which are screened to remove ceased collisions. Afterwards, we only need to check collisions that involve dynamic objects. In other words, all static-static pairs can be safely pruned from testing. Finally, we label objects with speed below/above 0.05, in any direction, as static/dynamic, respectively. Empirically, we found that the incremental detection is faster than the plain algorithm whenever more than half of the objects are static. To have the best of both worlds regarding temporal optimizations and the lack of thereof, we alternate between both approaches in run-time based simply on the number of static objects labeled on the current frame.

In the full text of the M.Sc. dissertation, detailed analyses of the empirical observations used to guide the design of our solution are presented, as well as the technical and implementation details of each algorithm used.

V. Tests and Results

All implemented algorithms were tested and their average time per frame from one thousand to one million objects, as presented numerically in Table II, and graphically in Figure 2. For brevity, we show the results for same-sized objects only and the worst and best variants of each family, when more than one algorithm is available. To scope our analysis to the best solutions available, we chose variants of each family, when more than one algorithm is available. Table I: Bundled algorithms within the Broadmark system. Time complexity derived for the uniformly distributed case [11].

VI. Conclusion and Future Work

We are confident that this work may improve knowledge in the broad phase collision detection area.

In the M.Sc. dissertation [11] we answer the, until then, open question “is it possible to have a solution that is both general and scalable for the broad phase collision detection problem?”.

In addition, we contribute to the advancement of the field by releasing the Broadmark methodology as an open-source tool for the community.

With these initiatives, this work may have many implications for research into the collision detection area. More specifically, we hope for new research to be developed using the framework and that, in turn, for their results to be more easily reproducible and validated by others, increasing both the significance and impact of each individual contribution. Moreover, with Broadmark, we hope to contribute with lowering the barrier of entry for new researchers to the field, to raise the quality of following comparatives and to encourage the research for novel, general-purpose, solutions.

Finally, in the near future, we plan to investigate novel parallel CPU and GPU solutions while retaining the joint focus on generality and speed, and to further develop the Broadmark framework to include new algorithms and scenes, as well as to support other related tasks, such as collision queries and continuous detection. Related to the topic, we are interested in applying deep learning models to collision detection research, on topics such as cloth and volume deformations.

VII. Awards and Publications

The two main results of this work, i.e., the new general and scalable solution for broad phase collision detection and the new methodology for the comparative analysis of algorithms named Broadmark, have been individually published in the Computer Graphics Forum (CGF) journal [4], [12] (Qualis A1), and come from the exclusive effort of the student and advisor, not being part of an overarching project.

In addition, during the M.Sc, two other journal publications related to this work have been published in the Computers in Entertainment [24] (Qualis B1) and ACM Entertainment Computing [25] (Qualis B1) journals.

Moreover, the authors have been formally invited by the program Co-Chairs of the ACM SIGGRAPH / Eurographics Symposium on Computer Animation (SCA’19) (Google Scholar h5-index 20) to present the results published on the CGF journal [4] as a journal first entry at UCLA – University of California, in Los Angeles, USA, testifying the quality and relevance of the research to the scientific community.
The SCA’19 program schedule and invitation letter can be found, respectively, at [https://sca2019.kaist.ac.kr/wordpress/program/] and [https://1drv.ms/b/s!Aq35PBOZWmsjhsUyoq9eevMYa5pD-w].

In parallel, still in 2019, the field of artificial intelligence started to be studied as a complementary tool. This study led to other two scientific contributions: the presentation of a tutorial on the Conference on Graphics, Patterns and Images (SIBGRAPI’19) [26], and a First Place Award-Winning full paper in the Computing Track of the Brazilian Symposium on Games and Digital Entertainment (SBGAMES’19) [27].

Finally, this work has also recently won a prestigious award from the XXXIII Concurso de Teses e Dissertações (CTD) of the XL Congresso da Sociedade Brasileira de Computação (CSBC): It is among the Top 10 finalists chosen from 73 M.Sc. dissertations in the Computing area in Brazil, concluded in 2019. The 3 winners will be announced next November, during the CSBC conference, after online presentation of the 10 finalist works.

REFERENCES

Figure 2: Scalability analysis up to one million objects.


