Modelling Sperm Behaviour in a 3D Environment

M. Burkitt
Department of Computer Science
University of Sheffield
Sheffield, UK
m.burkitt@sheffield.ac.uk

D. Walker
Department of Computer Science
University of Sheffield
Sheffield, UK
d.c.walker@sheffield.ac.uk

D.M. Romano
Department of Computer Science
University of Sheffield
Sheffield, UK
d.romano@sheffield.ac.uk

A. Fazeli
Academic Unit of Reproductive and Developmental Medicine
University of Sheffield
Sheffield, UK
a.fazeli@sheffield.ac.uk

ABSTRACT

The processes used by mammalian sperm to find the egg in the female reproductive tract are highly complex and poorly understood. Due to ethical and practical limitations, observing and measuring the behaviour of sperm within a live animal is difficult if not impossible without influencing their behaviour. One way to help understand the processes involved is through the use of computational modelling. The methodology used to construct the first agent based model of sperm movement within a 3D model of the mammalian oviduct is presented. The different processes represented within the model and the implementation of those processes is described. The simulation runtime is significantly reduced using collision detection optimisation and by harnessing the parallel processing power of the GPU to make concurrent simulation replicates. Validation of the model and the potential uses of the model once fully validated are described.

Categories and Subject Descriptors

I.6.5 [**Simulation and Modelling**]: Model Development—*Modeling methodologies*; J.3 [**Life and Medical Sciences**]: [Biology and genetics]

1. INTRODUCTION

Understanding the mechanisms that influence the movement of sperm within the mammalian oviduct, which is called the fallopian tube in humans, is essential for identifying the causes of infertility, tracking the progression of disease and developing new methods of contraception and assisted conception. The combined influence of different mechanisms results in the regulation of sperm progression along the oviductal tube to the site of fertilisation. Each of these mechanisms contains multiple factors which could potentially influence the likelihood of successful fertilisation. However, the relative significance of each factor is not fully understood. The resulting complexity, combined with the limitations in observing the reproductive system using biological experiments, make this difficult to investigate. One way of simplifying this complexity is through computational modelling. Here we present the methodology used to construct the first agent based computational model of sperm movement within the oviduct.

The main aim of this initial model is to investigate how the combination of basic sperm navigation mechanisms affect the distribution of sperm over time within an accurately scaled 3D model of the oviduct. By modelling the behaviour of sperm individually, we can identify the impact that variations in the parameters of each mechanism have on the temporal distribution of sperm. This can help us to understand how the combination of all these processes results in the regulation of sperm passage to the site of fertilisation. This paper focusses on the techniques used and technical aspects of model creation.

First a conceptual model of sperm behaviour is established based on data from literature. The properties of the biological system are described, and the implementation of each aspect is explained. The model is implemented using FLAME GPU [23], which is an agent based modelling framework which runs on NVIDIA graphics cards using Compute Unified Device Architecture for C (CUDA). The techniques used to handle interactions between the sperm agents and their 3D environment are described, as well as the optimisation of collision detection. Management of the massive amount of data generated by the model, and techniques for data analysis and validation against literature sources are also discussed.

2. BACKGROUND

The mammalian oviduct has a convoluted tubular structure, which connects the ovaries to the uterus. The inside of the oviduct contains soft tissue folds which change in complexity along the length. The oviduct can be split into several

CMSB '11, September 21-23, 2011 Paris, FR

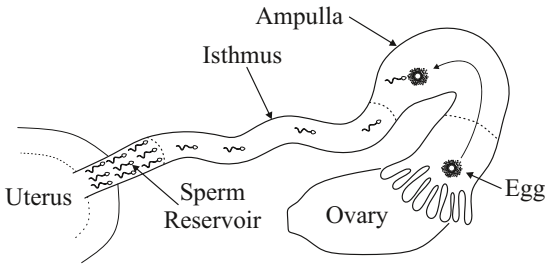

Figure 1: Shows the main regions of the oviduct, along with an overview of the progress of sperm and egg within the system.

sections, each with a different internal structure and function. The main components of the oviduct, as shown in Figure 1, are the isthmus and the ampulla. Sperm, which are the male reproductive cells, enter the oviduct from the uterus and meet and fertilise the female reproductive cells, commonly referred to as the oocyte or egg, in the ampulla.

Sperm are autonomous individuals which swim through liquid in the oviduct using a tail-like flagellum for propulsion. The way the flagellum beats can result in different movement patterns, which typically present themselves as either progressive movement, which moves the sperm forward, or a more vigorous non-progressive movement, which moves the sperm in an erratic or circular manner [9].

Even though several million sperm are initially deposited, by the time they reach the oviduct, they have been through several complex processes and only a small subset of the original population remain [27]. Damaged, weak or malformed sperm are prevented from passing into the oviduct in many species [10]. This means that the majority of sperm within the oviduct, which can range from a few hundred to a few thousand depending on the species, are capable of fertilisation.

Upon entering the oviduct, sperm in many species are stored in a 'sperm reservoir' at the most caudal part of the isthmus by attaching to the epithelial cells that line the wall of the oviduct [27]. While stored in this nutrient rich environment, sperm remain in an uncapacitated (inactive) state [27]. Over time, sperm gradually become capacitated (activated) and detach from the epithelium [27]. Capacitated sperm move through the oviduct, periodically re-attaching to the epithelial wall and then detaching and moving on. Sperm do not attach every time they collide with the oviduct, and may instead swim off in a random direction [9]. It has been observed that sperm periodically switch between progressive and non-progressive movement, and that non-progressive movement patterns may be required for detachment from the epithelial wall [9]. Eventually, sperm find and attach to the oocyte and the process of fertilisation begins.

The exact mechanisms controlling this process are not fully known, and several theories such as individual sperm movement; the physical structure of the oviductal environment; the influence of fluid dynamics on sperm transport; sperm navigation through thermal and chemical attraction and localised interactions of sperm with oviductal secretions at a molecular level are being investigated by researchers [25, 27, 30]. A combination of these behaviours results in the gradual passage of sperm from the sperm reservoir to the site of fertilisation. This regulated progression is important because if multiple sperm reach the egg at the same time, then polyspermy can occur. This is where a single egg is fertilised by more than one sperm, resulting in abnormal development and early termination. Under normal conditions, the slow stream of sperm results in at most a 1:1 ratio between sperm and eggs at the site of fertilisation until all eggs have been fertilised [30]. It is important to note that different mechanisms may have a more or less prominent role in different species.

It is clear that when modelling sperm populations, the properties and behaviour of individuals are important. The chosen approach was to model sperm at a cellular level using Agent Based Modelling (ABM). ABM is a technique that is used to model the individual behaviour and interactions of entities within a system [3]. The simple behaviour of individuals and how they interact with their environment is explicitly modelled, and the complex population level behaviour is an emergent property of the simulation. ABM techniques are increasingly being used to model the behaviour of individual cells. Walker et al. [28] used ABM techniques to model individual epithelial cells in combination with *in vitro* modelling to investigate wound healing in epithelial cell monolayers. The immunity response of T cells when interacting with dendritic cells within lymph nodes was modelled by Bogle and Dunbar [2].

When considering interactions between sperm agents and a large 3D environment, collision detection algorithms are needed. A naïve collision detection implementation will perform intersection tests between a moving object and every triangle in the environment, resulting in $O(n^2)$ triangle-intersection tests, where n is the number of triangles. The majority of collision detection algorithms aim to reduce the number of intersection tests as much as possible.

One method commonly used to reduce the number of intersection tests is scene partitioning. The space in which objects exist within the 3D environment is split into sections and only objects within each section are tested for intersections. Various partition methods exist, such as Octrees and Binary Scene Partitioning (BSP), which group the objects together in different ways based on their spatial position [29].

Another efficient method is through the use of bounding volumes. A complex object is wrapped in a simple object and collision tests are performed between the bounding objects to eliminate pairs of objects that are too far apart to collide. Where the bounding volumes of two objects do collide, a per triangle intersection test is then performed between the underlying objects to determine if any collision has taken place [29]. Although many collision detection algorithms are implemented on the CPU, more recent versions have been partially or completely implemented on the GPU, with significant performance gains [17, 20].

Previous computational models of sperm have been created to investigate their movement within artificial environments. These include computational modelling of sperm behaviour in a microfluidic sperm sorter [15]; using mathematical formulas from hydrodynamics theory to investigate the phenomenon of sperm accumulation near surfaces [26]; and investigating the crawling behaviour of Nematode sperm at a mechanochemical level using finite element modelling [4]. However, none of these models attempt to investigate the behaviour of sperm within their natural environment.

3. MODELLING SPERM

The agent model is presented using the Overview, Design concepts and Details (ODD) protocol proposed by Grimm et al.[13]. The model described here is largely generic, and can be easily modified to simulate the behaviour of sperm in different mammals. In order to demonstrate the capabilities of the model, data from a single species was used. Due to the relative abundance of available data, the mouse was chosen as the target species for investigation, and the value of all parameters described are based on the mouse where possible.

3.1 Purpose

The purpose of the model is to investigate the relative significance that proposed mechanisms related to sperm movement in the female reproductive tract have on the spatial population distribution of sperm and the average amount of time taken for sperm to reach the oocyte.

3.2 State Variables and Scales

The main individual within the model is the sperm agent, which moves within the 3D environment of the oviduct. Sperm are characterised by the state variables: id, 3D position and direction (encoded as a 4x4 transformation matrix), movement state (progressive or non-progressive), capacitation state (uncapacitated, capacitated or dead), collision state (free, touching epithelium, attached to epithelium or attached to oocyte), movement state and capacitation state timers (amount of time remaining in the current state) and oviductal slice (current slice of the oviductal tube). Oocytes are represented by a static position in 3D space (3 component vector).

The size of sperm and oocyte agents are based on statistical studies and measurements from images [11, 16]. The speed and movement characteristics of sperm are obtained from literature reporting Computer Assisted Semen Analysis (CASA) measurements of sperm movement [12]. The table in Figure 1 shows the global model parameters.

The environment is an accurately scaled 3D model of a mouse oviduct, which appears as a convoluted tube approximately $17,000\mu m$ in length, and is subdivided into sequential segments. This complex 3D structure was created using techniques described in [6, 7]. Histology images from mouse oviducts were processed using semi-automated techniques to identify the structure of the 3D tube and to obtain measurements of the complexity of the internal soft tissue folds. These measurements were used to parameterise a particle based construction algorithm. The parameters were further modified based on measurements and qualitative descriptions from literature to take into account the expansion of the site of fertilisation due to fluid [30]. Multiple variants of the internal structure and multiple external bend structures were generated, resulting in nine unique models. One of the resulting models is shown in Figure 2.

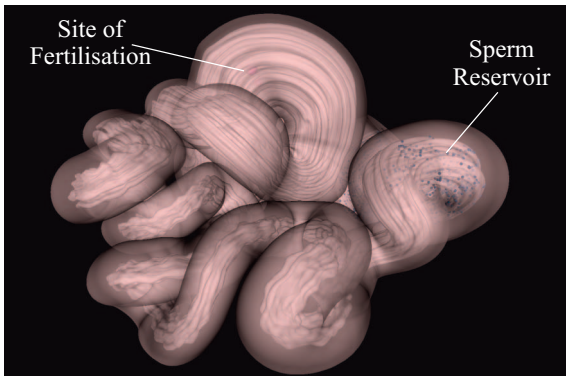

Figure 2: Shows one of the 3D oviduct models used for the simulations. The initial position of sperm agents within the sperm reservoir and the oocytes at the site of fertilisation are shown.

All position and distance values are in μm and the 3D environment is scaled with a mapping of 1:1 between μm and 3D units. The environment is limited to the extents of the tube of the oviduct model. The simulation time step represents 1 second of real time.

3.3 Process Overview and Scheduling

For sperm agents, several different process branches are followed based on the current state of the sperm. The main processes are: *Try to Capacitate*, *Move*, *Try to Attach to Epithelium*, *Try to Detach from Epithelium*, *Regulate Movement State* and *Regulate Capacitated Life*. The order in which the relative processes are called is shown in Figure 3. The *Try to Attach to Epithelium* process is called within the *Move* processes whenever a collision with the environment occurs. Collisions between sperm and oocyte are detected and cause the sperm to switch to the *attached to oocyte* state. All agents perform the same process at the same time in parallel.

3.4 Design Concepts

Emergence. The temporal evolution of the distribution of sperm within the environment is an emergent property of the system.

Interaction. Although many agent based simulations consider the interaction between individuals to be important, this simulation focuses on the interaction between the individuals and their environment. Therefore, interactions between sperm and the environment are modelled explicitly. When moving, if sperm collide with the environment, then they may either attach to the epithelium or turn and move off in another direction. Collisions between sperm and

Parameter	Type	Value	Source
Time in Progressive State	Time Range (s)	$10 - 60$	Estimated from observations [9]
Time in Non-Progressive State	Time Range (s)	$5 - 10$	Estimated from observations [9]
Attachment to Epithelial Threshold	Probability	0.1	Literature guided calibration [9]
Detachment from Epithelial Threshold	Probability	1.0	Estimated from observations [9]
Capacitation Threshold	Probability	9.26×10^{-5}	Literature guided calibration [5]
Oocyte Fertility Start Time	Time (s)	16,200	Estimated from data [30]
Capacitated Sperm Life	Time (s)	7200	Estimated [1]
Sperm Radius	Size (μm)	1.6	Average from measurements [11]
Oocyte Radius	Size (μm)	36.65	Measurements from images [16]
Progressive Velocity	Velocity ($\mu m s^{-1}$)	146.9	From CASA measurements [12]
Non-Progressive Velocity	Velocity ($\mu m s^{-1}$)	73.3	From CASA measurements [12]
Sperm Population	Count (n)	$1,536$	Average from data [14]
No. of Oocyte	Count (n)	4	Average from data [5]
Max. Reflection Angle	Angle (deg)	60	Literature guided calibration [9]
Max. Non-Progressive Rotation Angle	Angle (deg)	60	Literature guided calibration [12]
Max. Detachment Rotation Angle	Angle (deg)	180	Estimated from observations [9]
No. of Movement Steps	Count (n)	4	Estimated

Table 1: Shows the global parameters used in the model.

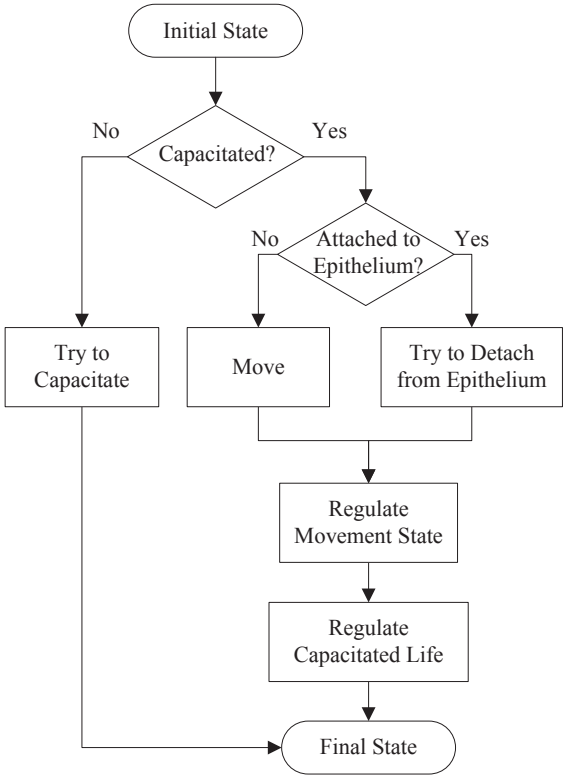

Figure 3: Shows the processes followed by each sperm agent every iteration.

oocyte are detected, causing the sperm to attach to the oocyte and change its state. Sperm to sperm interactions are not modelled, and sperm effectively pass through each other. Due to the small number of sperm which are capacitated at any one time and the large size of the oviductal environment in relation to the sperm agents, interactions between individual sperm are unlikely to have a significant impact on the progression of sperm.

Stochasticity. Stochasticity is used in many of the processes within this system. The initial position and direction of each sperm is determined stochastically. The processes *Try to Capacitate* and *Try to Attach to Epithelium* are both based on probabilities. The direction that sperm swim when reflecting off and detaching from the epithelial and swimming non-progressively is stochastically determined within a constrained range.

Observations. As the model runs, the state variables of all agents are recorded at every time step. This is used to observe how the relative distribution of sperm within the environment changes with respect to time. The time taken for sperm to reach the oocyte, and the number of sperm meeting the oocyte within the simulation time and the number of oocyte with sperm attached are also used as measures of the system.

3.5 Initialisation

The sperm agents were initially distributed randomly along the wall of the lowest segment of the isthmus in the oviduct model [30]. The initial direction was randomly determined within a 60° cone relative to the direction of the tube. This ensures that all sperm were initially pointing in the correct direction, as would be the case if they had just entered the tube from the uterus. All sperm were initially uncapacitated [27].

3.6 Input

The main inputs into the simulation are the 3D environment; the initial position, direction, state and number of sperm; and the position and number of oocyte. The sperm input parameters are based on literature or estimated where data is unavailable.

3.7 Submodels

Try to Capacitate. The *Try to Capacitate* process determines if currently uncapacitated sperm should become capacitated at the current iteration. It is encoded as a simple probability, the value of which was estimated based on the assumption that the majority of uncapacitated sperm will have capacitated before the end of their lifetime, which is estimated to be at most 12-14 hours, with a significant drop in fertility observed after 4-6 hours [21]. The *Capacitation Threshold*, which is the probability (P) that a single sperm would become capacitated at any second was calculated in relation to the number of hours (n) it should take for all sperm to become capacitated as:

$$P = \frac{1}{(n \times 3600)}$$

with $3,600$ being the number of seconds in an hour. Once capacitated, the sperm have their *capacitation state timer* memory variable set to the *Capacitated Sperm Life* value, which determines how long sperm can live for once capacitated. n was set to 3 after model calibration.

Move. Capacitated sperm which are not attached to the epithelium move in different ways depending on their movement state. During a single iteration, progressive sperm move forward at a fixed rate, determined by the *Progressive Velocity*. Non-progressive sperm movement is more erratic, so a instead of moving forward, the sperm agents turn randomly based on the *Max Non-Progressive Rotation Angle* and then move in this direction at a the *Non-Progressive Velocity*.

The potential new position is calculated, then any collisions with oocytes or the environment are identified. Collisions with oocyte are determined if the sphere representing the sperm, with radius *Sperm Radius* intersects the sphere representing the oocyte, with radius *Oocyte Radius*, during sperm movement. If a collision with an oocyte is made, then the sperm are moved to a *attached to oocyte* state, their position is set to the collision point and they take no more part in the simulation. Collisions with oocytes are only permitted once the oocyte has had sufficient time to mature, which is determined by the *Oocyte Fertility Start Time* property. If a collision occurs before the oocyte is fully matured, then the sperm will pass straight through. If a collision with the environment is made by sperm with either progressive or non-progressive movement, then the *Try to Attach to Epithelium* process is followed. If the sperm does not attach to the epithelium, then a new direction is calculated using the angle of incidence, the surface normal of the collision point, and the *Max. Reflection Angle* property.

In reality, sperm move with a series of flagellar beats, and can potentially collide with the environment and change direction multiple times during a single second. To reflect this, the movement of progressive sperm was split into a series of sub-steps, based on the *No. of Movement Steps* property. Every iteration, the *Move* process is repeated multiple times (*No. of Movement Steps*), with each agent moving a fraction of the total (*Progressive Velocity*). This means that sperm can move, collide with the environment and change direction multiple times each iteration.

Try to Attach to Epithelium. If progressive sperm collide with the oviduct wall, then they may or may not attach to the surface [9]. This parameter (*Attachment to Epithelial Threshold*) is encoded as a probability, the value of which is uncertain. The probability of attachment was systematically investigated, and was set to 0.1 after model calibration. Sperm which do attach to the epithelium are switched to the *attached to epithelium* state.

Try to Detach from Epithelium. Sperm which are attached to the oviduct epithelial usually detach when they switch to non-progressive movement [9]. Based on the qualitative description, sperm are assumed to detach whenever they switch to non-progressive movement. When sperm detach from the epithelium, a new direction is calculated using the angle of incidence, the surface normal of the attachment point, and the *Max. Detachment Rotation Angle* property.

Regulate Movement State. It has been observed that sperm typically have 5 - 10 second bursts of non-progressive movement, followed by 10 - 60 seconds of progressive movement [9]. To represent this, uniform distributions based on the ranges for progressive movement (*Time in Progressive State*) and non-progressive movement (*Time in Non-Progressive State*) were used to calculate the amount of time a single sperm should remain in its current movement state.

Whenever a sperm switches movement state, then the amount of time it should remain in the new state is randomly selected from the corresponding uniform distribution and stored in the *movement state timer* memory variable. Each iteration, this value is reduced by one and when it reaches zero, the movement state is switched and a new timer variable is calculated.

Regulate Capacitated Life. The lifetime for capacitated sperm which was set during the *Try to Capacitate* process is reduced by one each iteration. When this value reaches zero, The sperm state is set to *dead*, and the agent takes no more part in the simulation.

4. MODEL IMPLEMENTATION

This section describes the computational realisation of the conceptual model of sperm behaviour.

4.1 Agent Model

As the model interacts with a complex 3D environment, a large number of mathematical calculations were required every iteration to handle collision detection and resolution. To handle this requirement, high performance computing techniques were used to implement the model. Flexible Large-scale Agent Modelling Environment (FLAME) is an agent based modelling framework which allows agent based models

to be created and run on high performance grid computers without the need to write specialist code [8].

A recent variant of the FLAME framework called FLAME GPU [22] allows these models to be run using a consumer NVIDIA graphics card using CUDA [18]. Graphics cards (GPUs) are primarily designed for displaying 3D images on the screen at a high frame rate. They are highly optimised for performing a large number of fast mathematical calculations in parallel. In recent years, the GPU has been used for scientific computing, as the fast, parallel nature of the hardware is ideal for performing certain types of computational simulation [18]. CUDA is a set of extensions to the C programming language which makes it simpler to access the memory and run functions, which are called kernels, on the GPU. The massively parallel capabilities of graphics cards and their optimisation for 3D mathematical calculations made this an ideal choice for this type of model. FLAME has been successfully used to model epithelial cell behaviour, with FLAME GPU demonstrating significant performance gains over the original [23].

To create a model using FLAME GPU, an XML based definition file is created which describes each agent and their properties, how agents interact with each other, and provides a set of function declarations. The implementation of each function is then written in CUDA, with each function forming a CUDA kernel. Code for managing state transitions, agent lists, memory allocation, data persistence and kernel execution are all automatically generated from the XML template. This greatly simplifies model development, allowing the model developer to focus solely on the model without worrying about the complexity of parallel programming.

4.2 Sperm to Oviduct Interactions

Interactions between sperm and the oviductal environment can have a significant impact on sperm navigation and behaviour. In order to model these interactions computationally, an algorithm was devised to efficiently detect collisions. Many collision detection strategies combine multiple different techniques, which are selected depending on the shape, size, complexity and behaviour of objects within a 3D environment.

As a basis for the collision detection, a Sphere Sweeping technique [24], modified to run in CUDA, was used. A sphere representing each agent is projected along its intended path and the fraction along the path where the sphere intersects with each triangle is detected. For each triangle, the current sphere position and direction are used to determine if the sphere can collide with the plane that the triangle lies on. If an intersection is possible, then the position and direction are converted into 2D space, relative to the triangle plane. The circle formed by the projection of the sphere in 2D is then tested for intersection with the triangle [24]. As the triangle-intersection test is computationally expensive, the number of intersection tests performed each iteration should be reduces as much as possible. Therefore, a combination of scene partitioning and bounding volumes were used to minimise the number of intersection tests.

As the 3D environment is essentially a long tube, it can be split into sequential slices. Each sperm agent tracks its current position along the tube. During movement, each agent determines the slices with which it could potentially collide. The line formed between the current position and the potential new position, which is calculated using the direction vector and the movement distance, is used to perform a series of intersection tests with the plane marking the edge of each slice. From the starting position, slices are sequentially tested in-front of the current slice and then behind the current slice. The process stops if the plane from a slice is not intersected. This is illustrated in Figure 4. The initial slice range is set to one slice before and one slice after the current slice. Intersections with the slice planes are progressively checked from (a) to (g). The movement path does not intersection with slice (d) so the algorithm interrupts the process and marks the outer boundary as slice (d).

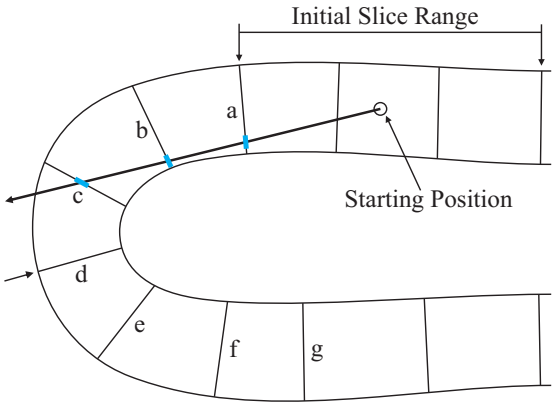

Figure 4: Shows how the potential slices are identified. The initial slice range (3 slices) is progressively extended as intersections with the corresponding slice planes are identified. The algorithm continues until slice plane (d) is processed where no intersection occurs. The final slice range is 6.

Reducing the number of triangles by slice identification significantly reduced the number of triangle-intersection tests, but this was further reduced by wrapping each triangle in a bounding sphere. Before performing the triangle intersection test, the distance between the start point and the triangle bounding sphere was tested. If it was larger than the maximum moving distance, then the triangle intersection test was not needed. If it was within the maximum moving distance, and could therefore potentially collide, then a line-sphere intersection test was performed. If the line intersected the sphere, then finally the triangle-intersection test was performed.

3D models are typically collections of vertices within 3D space, with triangle definitions linking vertices, resulting in a set of surfaces. As the collision detection algorithm performs intersections in 2D space relative to each triangle, it was inefficient to store the model in a traditional manner. For optimisation with the collision detection algorithm, instead of storing the triangle and vertex positions in 3D space,

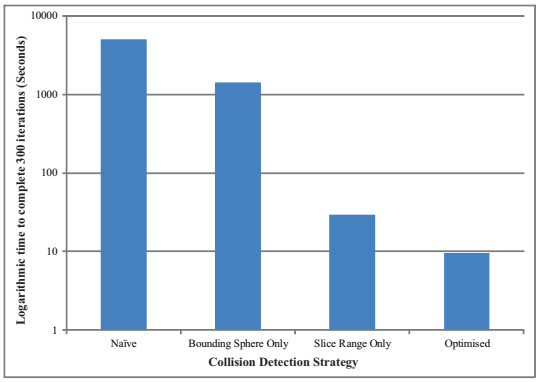

Figure 5: Shows the time taken to complete 300 iterations using different collision detection strategies. The scale is logarithmic due to the exponential difference in time between the naïve and optimised approaches.

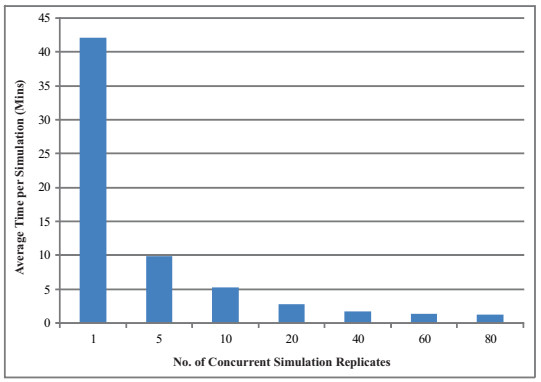

Figure 6: Shows the average time to complete a single simulation when multiple simulations are run concurrently.

the x and y position of the vertices forming each triangle in 2D space was pre-calculated and stored in the primary data structure. The collision detection algorithm uses the triangle plane (normalised vector + w offset), the vector perpendicular to the plane (U) and the cross product of the plane normal and U to convert between 3D space and 2D space local to the triangle. These values were also pre-calculated and stored in the primary data structure. This reduced the number of calculations required each iteration. A bounding sphere for each triangle was pre-calculated and used for the initial intersection tests. At the start of the simulation, the data structures containing the full definitions for the 3D environment were loaded onto the GPU as texture memory, which is optimised for read only access.

A set of simulations were run in order to test the significance of the collision detection optimisation. Each simulation contained $1,536$ agents, using one of the oviduct models which is made up of $648,930$ triangles split into $1,200$ slices. Each simulation ran for 300 iterations, representing a simulation time of five minutes. Figure 5 shows the speed of the naïve collision detection algorithm, along with the inclusion of scene partitioning, bounding spheres and the fully optimised solution which contained both. The naïve algorithm took over 83 minutes to complete 300 iterations, whereas the optimised algorithm took just over 9 seconds to complete the same number of iterations. Due to the massive difference in time, the scale is shown logarithmically for comparison. As a full simulation could potentially run for over 28,800 iterations, the value of collision detection optimisation becomes clear.

4.3 Running a simulation

To run a simulation, an XML file is created which defines the initial state of each agent in the system, and specifies the value of global constants. FLAME GPU is highly optimised for processing very large numbers of agents concurrently [22]. As this simulation only requires a relatively small number of agents, the model was modified to allow multiple simulations to run concurrently. Each agent was associated with a unique environment id, and only agents with a matching id could interact with each other. This means that multiple replicates of the simulation can be run concurrently. To demonstrate the benefit, the time taken to run a simulation with $1,536$ agents for $28,800$ iterations, with different numbers of concurrent replicates was recorded.

Figure 6 shows the average time for each simulation to complete when different numbers of replicates are run concurrently. The average time is calculated as the total run time divided by the number of concurrent simulations. The results show that when 10 simulations are run concurrently, the average time per simulation is significantly lower than a single simulation. When the number of concurrent simulations reaches 40, then the maximum capacity of the GPU is achieved, and the relative performance increase is linear, with each simulation taking approximately one minute to complete.

All simulations were run using the following configuration:

- Windows 7 x64
- Intel Core 2 Quad Q6700 @2.66Ghz [4GB RAM]
- NVIDIA GeForce GTX 280 [1GB]

4.4 Data Management & Analysis

Typically, agent based simulations record the position and state information for each agent every iteration. When run for a simulated 8 hour period ($28,800$ iterations) with $1,536$ agents and multiple replicates, this could potentially result in a massive amount of data.

The FLAME GPU framework generates data in XML format, and the initial 0.XML file for a 40 replicate simulation was 47MB, resulting in a potential total results set of

1.3TB per simulation run, which is clearly unmanageable. The framework was modified to store the simulation data in binary format, with an initial header file describing the structure of the data, and then the iteration records following sequentially. The location of data for each iteration was stored in a lookup table in the file header. This significantly reduced the potential final simulation data size down to 130GB. The framework was further modified to only export data at specific intervals, in this case every 1,800 iterations, representing every 30 mins. As the observational data in biological studies is usually oriented around hourly observational measurements, this provided sufficient data for the analysis. It also had the advantage that all the data required for statistically significant analysis of a single simulation run to be performed was stored in a single output file, with a size of around 40 MB. The data generated by the framework was aggregated to store the number of sperm per oviduct slice with respect to time. Each slice was mapped onto a corresponding region of the oviduct, such as the isthmus or the ampulla, so the number of sperm in each oviductal region at specific time points could easily be established.

5. POTENTIAL FOR FUTURE WORK

The final aim of the model is to use it for making predictions about the likelihood of a specific population of sperm reaching the oviduct, under a given set of conditions. For example, based on the sperm count and motility parameters obtained from CASA analysis, the likelihood of a specific sample of sperm reaching the oocyte and resulting in successful fertilisation can be established. Abnormal chemical or structural changes within the oviduct could also be represented. This would have uses not only for assessing the need for assisted conception in humans, but also for insemination and breeding in livestock.

Before a model can be used for these type of predictions, it needs to be fully validated. To fully validate this type of computational model, complementary biological experiments are required to investigate some of the mechanisms where data is lacking in literature. In the absence of complementary experimental work, the presented model of sperm behaviour can only be partially validated.

At the lowest level, face validation of a computational model can be obtained if the reasonableness of assumptions can be established, and the resulting simulation appears correct. This is often a very subjective process, based on discussions with experts on the system being modelled [19]. One approach proposed by Klügl [19] is to show animations of the results, showing the group behaviour and individual behaviour, to experts in the field and obtain verbal agreement that the observed behaviour is plausible. This can further be enhanced if the patterns of behaviour in the results generated by the model are statistically similar to patterns observed in the real system [19].

This can be complemented by sensitivity analysis, which identifies the significance of small changes in individual parameters with respect to the model output. This is especially useful for parameters where the value is estimated or was determined through model calibration, as it can help to identify the potential significance of errors in the estimated value. Where the value of a parameter is completely unknown, the complete range of values for a parameter can be tested, starting at the minimum value and increasing in regular intervals until the maximum value. This allows the total potential influence of the parameter to be investigated.

For this model, calibration, sensitivity analysis and uncertainty analysis have been performed, and model validation is in progress. Visual animations of the simulation have been presented to experts in the field, who considered the behaviour to appear reasonable, giving the model face validation. Preliminary analysis of the distribution of sperm within the oviduct over time appears to be similar to that described in literature, although this is still to be demonstrated statistically. The model is grounded in reality, with accurate spatial and temporal scales used throughout. The results of this will be published at a later date once completed.

6. CONCLUSIONS

The methodology used to construct the first agent based model of sperm movement within the mammalian oviduct has been presented. The important behaviour described in literature has been represented, and how that behaviour was implemented in the model was described. The simulation runs within an accurately scaled 3D model of the oviduct, which has both the internal and external complexity represented. The model was parameterised using values derived from literature where available, and grounded in reality both spatially and temporally.

The model was implemented using FLAME GPU, which is a CUDA based agent modelling framework. Optimisation of sperm to oviduct interactions is described, and a 520 times performance increase was demonstrated when compared to the naïve implementation. Modifications to the model to allow independent simulation replicates to be concurrently performed is described, and a 33 times performance increase per simulation replicate was demonstrated. Management of the massive amount of data generated by each simulation is discussed.

Once validated, the model can be extended to investigate the potential influence of additional mechanisms such as chemotaxis and thermotaxis. The model could also be parameterised by data from other species, which will help to understand the relative significance of different mechanisms between species. Even with only partial validation, the model can still be used for generating hypotheses about how the system behaves, which could provide new directions for biological experimentation.

References

[1] J. M. Bedford. Significance of the need for sperm capacitation before fertilization in eutherian mammals. *Biology of Reproduction*, 28(1):108–120, Feb. 1983.

[2] G. Bogle and P. R. Dunbar. Agent-based simulation of t-cell activation and proliferation within a lymph node. *Immunology and Cell Biology*, 88(2):172–179, Nov. 2010.

[3] E. Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceed-*

ings of the National Academy of Sciences of the United States of America, 99(Suppl 3):7280–7287, May 2002.

[4] D. Bottino, A. Mogilner, T. Roberts, M. Stewart, and G. Oster. How nematode sperm crawl. *Journal of Cell Science*, 115(2):367–384, Jan. 2002.

[5] A. W. H. Braden and C. R. Austin. Fertilization of the mouse egg and the effect of delayed coitus and of Hot-Shock treatment. *Australian Journal of Biological Sciences*, 7(4):552–565, 1954.

[6] M. Burkitt, D. M. Romano, D. Walker, and A. Fazeli. 3D modelling of complex biological structures: The oviduct. In *EG UK Theory and Practice of Computer Graphics*, University of Sheffield, UK, Sept. 2010. Eurographics Association.

[7] M. Burkitt, D. Walker, D. M. Romano, and A. Fazeli. Constructing complex 3d biological environments from medical imaging using high performance computing. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 99(PrePrints), 2011.

[8] S. Coakley, R. Smallwood, and M. Holcombe. Using x-machines as a formal basis for describing agents in agent-based modelling. In *Proceedings of 2006 Spring Simulation Multiconference*, pages 33–40, Apr. 2006.

[9] R. P. Demott and S. S. Suarez. Hyperactivated sperm progress in the mouse oviduct. *Biology of Reproduction*, 46(5):779–785, May 1992.

[10] W. R. Dukelow and G. D. Riegle. Transport of gametes and survival of the ovum as functions of the oviduct. In A. D. Johnson and C. W. Foley, editors, *The oviduct and its functions*, chapter 7, pages 193–220. Academic Press, New York, 1972.

[11] R. C. Firman and L. W. Simmons. Sperm midpiece length predicts sperm swimming velocity in house mice. *Biology Letters*, 6(4):513–516, Aug. 2010.

[12] S. G. Goodson, Z. Zhang, J. K. Tsuruta, W. Wang, and D. A. O'Brien. Classification of mouse sperm motility patterns using an automated multiclass support vector machines model. *Biology of Reproduction*, In Press.

[13] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz, and G. Huse. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198(1-2):115–126, Sept. 2006.

[14] C. A. Hodges, M. R. Palmert, and M. L. Drumm. Infertility in females with cystic fibrosis is multifactorial: Evidence from mouse models. *Endocrinology*, 149(6):en.2007–1581–2797, Mar. 2008.

[15] T. Hyakutake, Y. Hashimoto, S. Yanase, K. Matsuura, and K. Naruse. Application of a numerical simulation to improve the separation efficiency of a sperm sorter. *Biomedical Microdevices*, 11(1):25–33, Feb. 2009.

[16] D.-H. Kim, D.-S. Ko, H.-C. Lee, H.-J. Lee, W.-I. Park, S. S. Kim, J.-K. Park, B.-C. Yang, S.-B. Park, W.-K. Chang, and H.-T. Lee. Comparison of Maturation, Fertilization, Development, and Gene Expression of Mouse Oocytes Grown In Vitro and In Vivo. *Journal of Assisted Reproduction and Genetics*, 21(7):233–240, July 2004.

[17] P. Kipfer. LCP algorithms for collision detection using CUDA. In H. Nguyen, editor, *GPU Gems 3*, chapter 33. Addison Wesley Professional, Aug. 2007.

[18] D. B. Kirk and W.-m. W. Hwu. *Programming Massively Parallel Processors: A Hands-on Approach*. Morgan Kaufmann, 1 edition, Feb. 2010.

[19] F. Klügl. A validation methodology for agent-based simulations. In *Proceedings of the 2008 ACM symposium on Applied computing*, SAC '08, pages 39–43, New York, NY, USA, 2008. ACM.

[20] S. Le Grand. Broad-Phase collision detection with CUDA. In H. Nguyen, editor, *GPU Gems 3*, chapter 32. Addison Wesley Professional, Aug. 2007.

[21] R. W. McGaughey, J. H. Marston, and M. C. Chang. Fertilizing life of mouse spermatozoa in the female tract. *Journal of Reproductive Fertility*, 16(1):147–150, June 1968.

[22] P. Richmond and D. Romano. Template-Driven Agent-Based modeling and simulation with CUDA. In W.-m. W. Hwu, editor, *GPU Computing Gems Emerald Edition*, Applications of GPU Computing Series, chapter 21, pages 313–324. Morgan Kaufmann, 1 edition, Feb. 2011.

[23] P. Richmond, D. Walker, S. Coakley, and D. Romano. High performance cellular level agent-based simulation with FLAME for the GPU. *Briefings in Bioinformatics*, 11(3):334–347, May 2010.

[24] J. Rouwé. Collision Detection with Swept Spheres and Ellipsoids, May 2003.

[25] A. Seytanoglu, A. S. Georgiou, E. Sostaric, P. F. Watson, W. V. Holt, and A. Fazeli. Oviductal Cell Proteome Alterations during the Reproductive Cycle in Pigs. *Journal of Proteome Research*, 7(7):2825–2833, July 2008.

[26] D. J. Smith, E. A. Gaffney, J. R. Blake, and J. C. Kirkman-Brown. Human sperm accumulation near surfaces: a simulation study. *Journal of Fluid Mechanics*, 621(-1):289–320, 2009.

[27] S. S. Suarez. Regulation of sperm storage and movement in the mammalian oviduct. *The International Journal of Developmental Biology*, 52(5-6):455–462, 2008.

[28] D. C. Walker, G. Hill, S. M. Wood, R. H. Smallwood, and J. Southgate. Agent-Based Computational Modeling of Wounded Epithelial Cell Monolayers. *Transactions on Nanobioscience*, 3(3):153–163, Sept. 2004.

[29] A. Watt and F. Policarpo. *3D Games: Real-Time Rendering and Software Technology, Volume 1 (With CD-ROM)*. Addison Wesley, Dec. 2000.

[30] L. Zamboni. Fertilization in the mouse. In Kamran and E. S. E. Hafez, editors, *Biology of Mammalian Fertilization and Implantation*. Charles C Thomas, Springfield, Illinois, USA, 1972.