

Evaluating the visual fidelity of physically based animations

Carol O’Sullivan
Trinity College Dublin*

John Dingliana
Trinity College Dublin*

Thanh Giang
Trinity College Dublin*

Mary K. Kaiser
NASA Ames Research Center†

Abstract

For many systems that produce physically based animations, plausibility rather than accuracy is acceptable. We consider the problem of evaluating the visual quality of animations in which physical parameters have been distorted or degraded, either unavoidably due to real-time frame-rate requirements, or intentionally for aesthetic reasons. To date, no generic means of evaluating or predicting the fidelity, either physical or visual, of the dynamic events occurring in an animation exists. As a first step towards providing such a metric, we present a set of psychophysical experiments that established some thresholds for human sensitivity to dynamic anomalies, including angular, momentum and spatio-temporal distortions applied to simple animations depicting the elastic collision of two rigid objects. In addition to finding significant acceptance thresholds for these distortions under varying conditions, we identified some interesting biases that indicate non-symmetric responses to these distortions (e.g., expansion of the angle between post-collision trajectories was preferred to contraction and increases in velocity were preferred to decreases). Based on these results, we derived a set of probability functions that can be used to evaluate the visual fidelity of a physically based simulation. To illustrate how our results could be used, two simple case studies of simulation levels of detail and constrained dynamics are presented.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—[Animation]

Keywords: animation, plausible simulation, collision handling, perceptual metrics, evaluation

1 Introduction

In computer graphics, there are many occasions where it is necessary to determine the quality of an image or animation. This could be to determine the extent to which a compression, geometric simplification, dynamic manipulation or computational lag will be perceptible, or to compare two algorithms in terms of both efficiency and visual quality. Quality metrics or heuristics are needed to achieve this, but quality is often difficult to define. When referring to images or sounds, the term *fidelity* is often used instead. In computer graphics, particular effort has been expended to develop perceptual metrics and heuristics to measure or predict the fidelity of images [Daly 1993; Ramasubramanian et al. 1999], rendering for

animations [Yee et al. 2001; Myszkowski et al. 2001] and simplified geometric models [Luebke and Hallen 2001; Watson et al. 2001]. Hodgins et al. [1998] explored the effect of model type on perception of human motion, while O’Sullivan and Dingliana [2001] qualitatively examined the factors that can influence the perception of anomalous collisions. However, no generic means of evaluating or predicting the fidelity of the dynamic events occurring in an animation exists to date.

The simulation community, who are typically concerned with large-scale applications for aviation and defense training purposes, have also been recently grappling with the concept of simulation fidelity [Roza et al. 2000; Longridge et al. 2001]. A recent report on Fidelity Definition and Metrics (FDM-ISG) attempts to specify the fidelity requirements of such simulations in a formal way [Gross 1999; Roza et al. 1999]. Simulation fidelity is defined as: ‘*The degree to which a model or simulation reproduces the state and behaviour of a real world object or perception of a real world object, feature, condition or standard in a measurable or perceivable manner*’. Fidelity requirements represent the level of realism the simulation must display in order to fulfil the user’s needs and objectives i.e., its *fitness for purpose*. System validation involves first measuring the current level of fidelity of a system against a *referent*, which can be the real world or some substitute for reality, followed by an evaluation of whether this fidelity level is acceptable for the intended purposes of the system. In this paper we are primarily concerned with the measurement problem and we will present a metric that is relative to a given referent (Daly’s Visible Differences Predictor (VDP) [1993] is another example of a relative metric, as it depends on a reference image to compare with, rather than providing an absolute measure of deviation from reality). We will describe also how such a referent can be derived for two simple case studies.

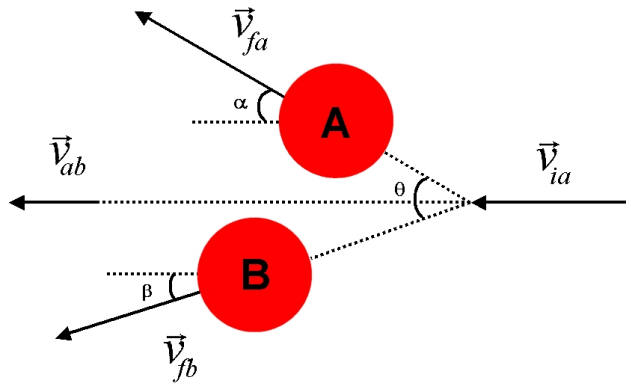
In Section 2, we provide the motivation for the development of a metric to assess the *visual fidelity* of physically based animations and consider some animation systems that produce dynamic distortions. In particular, Cheney and Forsyth’s [2000] probability-based definition of plausibility is discussed. In Section 3, inspired by an earlier study by Kaiser and Proffitt [1987], we present our psychophysical experiments that established some thresholds for the ability to perceive certain dynamic anomalies. Some interesting biases are identified e.g., expansion of the angle between the post-collision trajectories of objects was preferred to contraction; an increase in velocity following a collision was preferred to a decrease; higher acceptance thresholds were found for angular velocity than linear and object complexity also affected angular velocity perception; We use these results to fit a series of split Gaussian functions (see Figure 5) and discuss the construction of a simple but generalizable error metric by combining these terms with user-supplied weights. Finally, in Section 4, we demonstrate the application of this metric for two simple case studies: simulation levels of detail and constrained dynamics.

2 Background

The dynamics of a simulation can be distorted for a number of reasons and we will consider two main classes of anomalies, which we will refer to as *aesthetic* and *unavoidable* distortions respectively.

*e-mail: {Carol.OSullivan, John.Dingliana, Thanh.Giang}@cs.tcd.ie

†e-mail: mkaiser@mail.arc.nasa.gov



\vec{V}_{ia} is the initial velocity vector of object A (object B is initially stationary); \vec{V}_{fa} and \vec{V}_{fb} are the velocities after impact; \vec{V}_{ab} is the post-collision resultant velocity vector; α and β are the post-collision vector angles and $\theta = \alpha + \beta$.

Figure 1: Variables manipulated to study the perception of dynamic anomalies:

The former refers to systems where the ability exists to calculate correct dynamics, but for various reasons it is desirable to manipulate the collisions and trajectories in order to achieve a more aesthetically pleasing end-result. The latter refers to systems in which dynamic anomalies are the inevitable result of the limitations of the system.

2.1 Aesthetic Distortions

Popović et al. [2000] describe a system where rigid multi-body simulations can be manipulated by an animator, who can drag the bodies being simulated to any desired location. The system then computes the required physical parameters to direct the objects to that constrained position, by changing the initial positions and orientations of the objects or by adding new constraints to control properties such as surface normals or elasticity. A similar but simpler approach can be taken in games, where it may be necessary to keep certain paths clear for navigating agents, thus necessitating strategic ‘bounces’ that will direct any falling objects away from the critical area. Such techniques rely heavily on the subjective or aesthetic judgments of the human animator or developer, which may not always guarantee results that are visually plausible to a wide range of viewers.

Building on the work of Barzel et al. [1996], Cheney and Forsyth [2000] developed a system that exploits uncertainty in the world to generate a variety of animations to satisfy certain constraints. The user provides a simulator, the sources of uncertainty e.g., a function that quantifies the ‘goodness’ of an animation, and a set of constraints. The measure of goodness can be anything that the user desires e.g., it could be a person’s ability to detect an anomaly, or some other statistical measure, such as scanned real-world data. The simulator itself does not even need to be physically accurate, as plausibility is defined solely by the requirements of the user. A Markov Chain Monte Carlo (MCMC) process is used to sample the space of possible solutions and provide the user with more than one physically plausible simulation. This can be used in several ways, such as to ensure that something different happens in a game or training simulator every time it is run, thus making it more like the real world.

Cheney and Forsyth model uncertainty by defining functions that gives high probabilities for ‘good’ animations i.e., those that are most likely in the real-world and satisfy some user-defined constraints. For example, they present a simple animation of a ball bouncing and state that the probability should be high if all the

normal vectors used to generate the animation were close to vertical, and low otherwise. The direction of the normal vector for each collision is viewed as an independent random variable, distributed according to an unnormalized Gaussian probability distribution function with an arbitrarily selected standard deviation of 10° . The final probability $p(A)$ of an animation i.e., its plausibility, is calculated using the product of these probabilities and also takes the constraints to be satisfied into account.

While the distortion of the collision normal from the physically-accurate solution will certainly affect the perception of an anomalous collision, perception is not considered in the choice of standard deviation for the probability function. If the probability of a viewer *perceiving* an anomalous event was modelled instead, we could test if animations were visually plausible in addition to physically plausible. Therefore, a greater range of animations that satisfied the given constraints could be generated, which would be suitable for applications such as games where visual plausibility is more important than accuracy. Similarly, for an interactive manipulation system as described by Popović et al., a function that could assess the visual plausibility of the new animations generated would also be a useful tool for either automating the process or to help the animator in making his or her choice. The key to developing such a function lies in determining the standard deviation based on perceptual thresholds.

2.2 Unavoidable Distortions

In the systems described in Section 2.1, changes are made to the dynamics because the designer or programmer wants them to occur. In real-time animation and simulation, however, accuracy and realism must often be sacrificed in order to attain an acceptable level of interactivity. The inevitable, or unavoidable, outcome of such a trade-off is a simulation of lower quality than would be possible with a non-time critical approach. For faster rendering, geometry can be simplified and displayed at multiple Levels of Detail (LOD) in order to attain a target frame-rate and perceptual heuristics can be used to make this as imperceptible as possible [Funkhouser and Séquin 1993; Luebke and Hallen 2001]. Similarly, simpler physical models, called Simulation Levels of Detail (SLOD) [Carlson and Hodgins 1997; Cheney and Forsyth 1997; O’Brien et al. 2001; Brogan and Hodgins 2002] or simulation proxies [Cheney et al. 2001] can be used to reduce simulation times by applying simpler dynamics computations.

The criteria applied to choose these SLODs vary. Cheney and Forsyth [1997] describe a system where objects that are not currently in view are simulated with a lower level of detail. *Consistency* i.e., objects returning to view in a correct state and *Completeness* i.e., objects that should return because of their motions doing so, are effective criteria for SLOD selection in this type of system where only the behaviour of out-of-view entities is simplified. However, it is often necessary to have multiple levels of simulation detail in the visible scene also. Carlson and Hodgins [1997] describe a game where legged creatures are simulated with varying complexity, and state that the outcome of the game must be essentially unchanged by the use of SLODs and that the viewer’s perception of the motion must be the same. However, they found that “it is not possible to make definitive measurements of these criteria” and, as an approximation to a measure of visual plausibility, used a heuristic to measure the deviation in the outcome of the game by comparing the number of collisions with and without the accurate rigid body simulations.

O’Brien et al. [2001] generate a hierarchy of approximate motion models for a particle system and allow the SLODs to be selected at each time-step. They use the idea of Regions of Interest (ROI) - defined by the user, or by critical events in the simulation, such as collisions - to define the relative importance of groups of particles. Their primary interest is in the macroscopic behaviour of

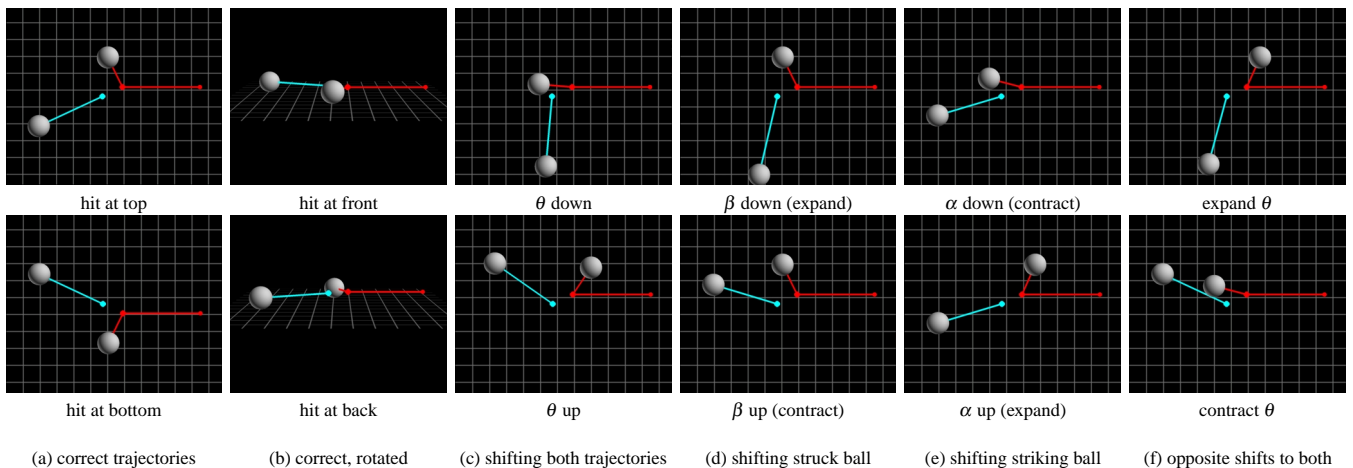


Figure 2: Angular distortions - The red and blue lines show the trajectories of the centers of the striking and struck balls respectively.

the system i.e., its visual appearance, and they try to improve this by avoiding situations in which larger deviations in motion paths occur. However, although they could produce data to quantify the speed-up factor achieved with their method, the lack of a definitive metric to evaluate visual quality - their stated chief criterion - meant that they could not provide empirical data to back up their claim that visual quality was not impaired. Instead, they had to rely on subjective judgments of the visual quality.

Simulation Level of Detail can also be achieved by using *collision proxies*. These are simplified volumes, sometimes hierarchically arranged, that are used instead of the rendering model in order to reduce the time taken for collision detection. For example, VRML and 3D Studio Max allow the user to specify a simpler volume, such as a bounding sphere, box or convex hull, to be used to determine any collisions with a more complex object. The processes of interpenetration determination and contact modelling are frequently the source of major bottlenecks in simulation (e.g., see [Mirtich 2000]), so the use of such proxies can lead to significant savings in terms of efficiency. Hubbard [1995] introduced the concept of *interruptible collision detection*, where a hierarchy of bounding spheres is used, each level forming a more accurate approximation to the actual object. When time runs out, the detection process simply returns a yes/no answer at the current level of accuracy. While Hubbard used a simple, non-physically based response, Dingliana and O’Sullivan [2000] described how information useful for contact modelling and response can be harvested during the collision detection process and then used to determine an approximate physically based response. The perceptual consequences of this degradation, such as perceivable gaps or inaccurate physical response, were then explored in [O’Sullivan and Dingliana 2001].

While spheres are very useful primitives for such interruptible algorithms, other hierarchies of bounding volumes such as Oriented Bounding Boxes (OBB) [Gottschalk et al. 1996], Discrete Oriented Polytopes (k-DOP) [Klosowski et al. 1998] or hierarchies of convex hulls [Ehmann and Lin 2001], may also prove to be good candidates for interruptible collision handling. Perhaps the improved fit and angularity of these volumes would offset any extra cost of processing them. To date these hierarchies have been mainly used to make early rejections or to quickly identify more complex regions of a model for accurate testing. However, without a metric that assesses simulation fidelity, including visual impact, objective comparisons between these different proxy models would be very difficult.

3 Experiments

We have seen that there is a need for a quantitative measure of how ‘good’ or plausible a physical simulation is. This measure is highly

dependent on the application and the aims of the developer. It could be a simple measure of deviation, considering uncertainty, as described by Chenney and Forsyth, or it could be a measure of fitness for purpose, as described by Roza et al. In many cases, such as those described in the previous section, the important factor is the perception of the animation by the human viewer. The development of a metric to evaluate this perception is not a trivial task. Based on the current state of psychophysics, we must build up such overall animation metrics from a multitude of narrowly-defined thresholds and tolerances. Simple, universal reductionist models that reliably predict perception are not yet known. Hence, we will explore in this section the factors that affect the perception of dynamic events, and describe a series of psychophysical experiments that aim to set some thresholds for the perception of simple elastic collisions between two rigid objects. These will then be used as a basis for developing the proposed metrics described in Section 3.3. We will discuss the outlook for future development of this metric in Section 5.

3.1 Previous results

Several studies have shown that people have certain weaknesses in detecting dynamic anomalies, that worsen as the dimensionality of the problems increase. Clement [1982] demonstrated that students’ intuitions concerning simple physics problems (i.e., their ‘naïve physics’) were highly inaccurate. Proffitt and Gilden [1989] showed that people are reasonably good at detecting dynamic anomalies when there is one dimension of kinetic information present i.e., what they call a particle motion, but are less competent with so-called extended body systems, where there are more dimensions of visual information e.g., both linear and angular velocity. In another study, Gilden and Proffitt [1989] showed that people used heuristics based on kinematic data, such as velocity after impact and the way that objects spun, to judge the relative mass of two objects involved in a collision.

O’Sullivan and Dingliana conducted a series of psychophysical experiments that examined the role of eccentricity (i.e., distance from the eye fixation position), separation, distractors, causality and accuracy of physical response on people’s ability to notice anomalous collisions. They found that erroneous collisions in the periphery of a viewer’s point of fixation are less likely to be detected, while collision anomalies that occur in the presence of increasing numbers of visually homogeneous distractors are also less noticeable. They also showed that adding a time delay between object contact and collision response reduced the perception of causality and thereby the plausibility of the collision event itself, as Albert Michotte [1963] also demonstrated. While their results suggest that

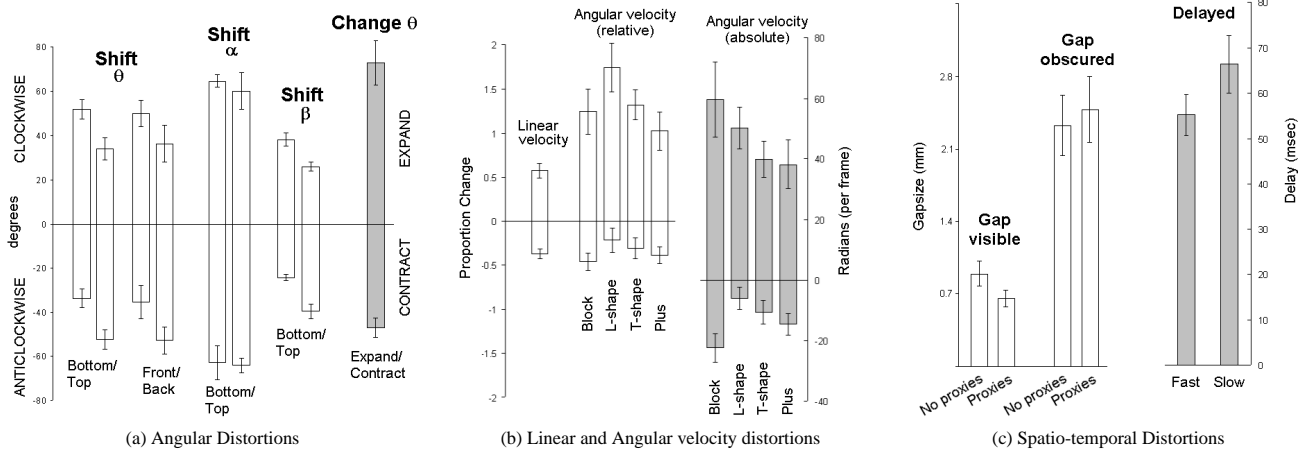


Figure 3: Results of psychophysical experiments

a random spin can often hide a bad collision and is sometimes even more plausible than the physically accurate response, this factor was not investigated in any further detail. Nor were the effects of the distorted post-collision trajectories and velocities on the ability to detect an anomaly investigated. We have seen in Section 2 that these and other types of distortions can occur and must be taken into account in any metric that attempts to measure visual plausibility.

We will look to a study by Kaiser and Proffitt [1987] to form the basis for our new experiments into human perception of dynamic anomalies. They presented a series of experiments that investigated the ability of people to discriminate between accurate and dynamically anomalous collisions. The methodology used in this study was a *randomly interleaved staircase* design [Cornsweet 1962; Levitt 1971]. The staircase (or up-down) method is an effective psychophysical technique for identifying thresholds, since it ensures that most of the trials are presented near the threshold for that particular observer. To establish an observer’s absolute threshold (i.e., the level at which a stimulus is detected 50% of the time), one starts with a level of 0 and increases the stimulus in large steps until a positive response is given. The level is then decreased (in smaller steps) until the observer gives a negative response, then increased until it results in another positive response. This ‘up-down’ process is repeated, sometimes with decreasing step sizes, until a pre-specified number of reversals is obtained (Kaiser and Proffitt used eight), at which point the process terminates.

This *ascending* staircase is complemented by a *descending* staircase, which starts at a clearly superthreshold level (i.e., the stimulus is glaringly obvious) and decreases until a negative response is given. It then reverses course, increasing the stimulus level until it evokes a positive response; then it proceeds with this up-down process until reaching the pre-specified number of reversals. To avoid observers anticipating the next stimulus (and hence biasing their response), trials from several staircases are *interleaved*; the trials then appear random to the observer. If observers respond systematically to the trials, the ascending and descending staircases should converge to the same threshold estimate. These threshold estimates are derived by taking the mid-point (i.e., the average) of the staircase’s last few reversal values (four is a typical number, and the one Kaiser and Proffitt used). In our studies, the ‘stimulus level’ was defined as the degree of distortion. We examined both *positive* (e.g., adding velocity at the point of collision) and *negative* (e.g., subtracting velocity) distortions.

A perceptual threshold, of course, is not an absolute point on the sensory continuum, but rather is stochastic. That is, a stimulus level is associated with a probability of detection. The function

commonly used in psychophysics to fit these response curves is a cumulative normal distribution, which is the integral of a normal (i.e., Gaussian) distribution. The general name for this family of S-shaped functions is an *ogive*; where the 50% ‘yes’ response rate intersects the ogive, the abscissa represents the absolute threshold stimulus intensity. This view of a threshold as a random variable fits very nicely with Chenney and Forsyth’s [2000] concept of plausibility as a probability function. They used a Gaussian function to estimate the probability that a deviation of the normal from the correct direction is acceptable. Later, we will derive a similar set of probability functions as a first step toward estimating the probability that viewers of an animation will perceive certain anomalies. The 50% intersection rate will be set using the results of our experiments.

In the original study by Kaiser and Proffitt, which we follow closely, a series of events were presented that depicted a stationary ball B being hit by a moving ball A - see Figure 1. Three different classes of dynamic distortion were carried out subsequent to collision:

1. Angular distortions: the post-collision trajectories of the objects were distorted by some angle. There were four different conditions for these distortions: the trajectories \vec{V}_{fa} and \vec{V}_{fb} of both disks were perturbed up or down i.e., θ remained constant but was shifted; the trajectory of the striking disk \vec{V}_{fa} or the struck disk \vec{V}_{fb} was distorted i.e., α or β was increased or decreased in isolation; the whole system was expanded or contracted i.e., θ was increased or decreased, distorting \vec{V}_{fa} and \vec{V}_{fb} while \vec{V}_{ab} remained constant.
2. Momentum distortions: the magnitude of the post-collision linear velocities \vec{V}_{fa} and \vec{V}_{fb} were proportionally increased or decreased.
3. Temporal distortions: the time at which the second ball was launched was distorted i.e., either there was a time delay induced at the point of collision, or the second ball started to move before the first ball hit it. (Note: the latter condition corresponds to the ‘gap’ effect explored by O’Sullivan and Dingliana).

In addition to distortions of the post-collision variables shown in Figure 1, we shall also consider anomalies in the angular velocity of objects after impact. Such distortions for collisions have

not been investigated before, although people’s ability to estimate angular velocity has been explored quite thoroughly. Barraza and Grzywacz [2002] present a survey of recent research in this area, and report on a set of experiments that investigated how humans could estimate how fast a rotational movement was. They asked people to compare the speed-of-rotation of pairs of stimuli composed of increasing numbers of dots. They found that increasing the number of dots improved discrimination performance. For a single dot moving in a circular motion, it is very difficult to know how fast it is rotating, as there are no visual clues to tell the brain that this is a rotation. When there are more dots, the spatial coincidence of several short trajectories provides a signal consistent with rotation to the brain, thus increasing the probability that the brain interprets the motion as rotational. In fact, several physiological studies have resulted in cortical cells sensitive to rotation and spiral motions being found (e.g., [Graziano et al. 1994]).

3.2 New Experiments

Inspired by Kaiser and Proffitt’s study (K&P), we ran a series of experiments to investigate the thresholds for dynamic distortion perceptibility. We followed their procedures closely to examine the effects of angular, momentum and *spatio-temporal* distortions¹. In addition, we investigated some new factors. K&P used only flat disks, whereas we used *three-dimensional stimuli*. We also used a *larger display* and examined the effect of different *viewpoints*. K&P recorded higher acceptance thresholds for the post-collision trajectories of their struck disk that were distorted downwards rather than upwards and attributed this to a preference for gravitational pull. However, in their study the first disk only hit the second disk from above, so we also varied the *hit-angle* to see if this had an effect. K&P did not test if *different linear velocities* affected thresholds, so we examined two different velocities for some of the angular and spatio-temporal distortion experiments. In the latter set of experiments, when the balls started to move before colliding (i.e., the early launch or gap effect), we used two different methods of computing the physics for these collisions, which we will refer to as being with or without *proxies*. Finally, we also conducted a new experiment to test anomalous *angular velocity* discrimination ability.

Kaiser and Proffitt repeated their experiments with and without friction, but found no significant differences, apart from an increased acceptance of momentum distortions in the absence of friction. We chose therefore to investigate only the non-frictional case in this study and will consider the role of friction in future work. In all experiments we used a staircase procedure as in the previous study.

3.2.1 Angular Distortions

We ran four experiments to find the acceptance thresholds for angular distortions of the post-collision trajectories of objects. These were run in a single block and 15 observers volunteered: 3 female and 11 male colleagues and students from the first author’s institution, aged 20 to 42, and one female child, aged 9. In these and all subsequent experiments, observers had normal or corrected-to-normal vision and were naïve as to the purpose of the experiments.

The stimuli were grey, diffuse-shaded spheres of diameter 4cm, on a 40 by 30cm display with a black background superimposed with a regular white grid². Participants sat at a distance of approximately 80cm from the screen, so the width of the display subtended

¹We refer to these distortions as spatio-temporal rather than just temporal, as the ability to detect if the balls started to move before colliding depends not only on temporal factors, but also on the visibility of the physical separation between them, which relates to space perception.

²Also used by Kaiser and Proffitt to minimize induced motion effects and to aid observers in trajectory angle estimation.

a visual angle of approximately 27°. When the event started, one ball was stationary and the other striking ball appeared from the right side of the screen, travelled horizontally, and collided with the stationary ball at an oblique angle (25° or -25°). A correct collision response was computed that assumed the balls to be of equal mass and, as in K&P, friction and gravity were ignored. Figure 2 shows the correct trajectories for both hit-angles, along with examples of distortions when the stationary ball was hit from above. The four experiments in this block were:

1. *Shifting θ* - the post-collision velocities of both balls were distorted in the same direction (see Figure 2(c)). For each condition, two staircases were started at the correct response, one ascending and one descending i.e., the trajectories were shifted upwards or downwards based on observers’ responses; another two started at the grossly distorted (up or down) positions and ascended or descended towards normal. We allowed the striking ball to have two initial velocities of 6 and 12cms per second and we also varied the hit-angle so that the stationary ball was hit at both the top and bottom. This gave rise to 16 staircases, which we randomly interleaved.
2. *Shifting θ , viewpoint rotated* - Identical to (1), with the viewpoint rotated horizontally away from the viewer by almost 90°, so the striking ball hit the stationary one in front or behind (see Figure 2(b)).
3. *Shifting α and β* - The post-collision trajectory of each ball was manipulated independently at two different hit angles, but at only one initial velocity (12cms/sec) and viewpoint (see Figure 2(d,e)). This also gave rise to 16 randomly interleaved staircases.
4. *Expanding θ* - In Experiment 4 both trajectories were manipulated simultaneously to cause θ to be expanded or contracted on each staircase (see Figure 2(f)), giving rise to 8 randomly interleaved staircases.

Each trial proceeded as follows: The grid and the stationary object appeared on the screen, the event was presented for a fixed number of frames (400 or 200 depending on velocity with 50% before and 50% after the collision). The screen was then cleared, and the observer had to press the left mouse button if the event appeared natural and the right button if it appeared distorted. The response was recorded, parameters were calculated for the next trial of that staircase (up or down a step of distortion), then another staircase was chosen at random and the program proceeded to the next trial of that staircase. The observers were instructed that gravity or friction did not influence the balls’ trajectories, that they should view the events as if they were happening in outer space and that the spheres were of equal mass. All staircases terminated after the eighth reversal i.e., the user had to switch 8 times from left button to right, or right to left.

Figure 3(a) shows the results for the four types of angular distortions, collapsed over velocity (which proved to be an insignificant factor). We took the average of the last four reversal values of each staircase as being an estimate of the acceptance threshold X_{50} i.e., the point at which the observer judged the distortion to be natural 50% of the time. Each type of distortion has two thresholds - a positive one and a negative one. Both have two estimates per subject, from the ascending and the descending staircases. We calculated the threshold estimates collapsed over staircase direction (no significant effects were observed for this factor in any of the experiments) by averaging over all observers. For all results reported in this paper, single factor ANOVAs (with 1 degree of freedom, unless otherwise stated) were conducted to determine the reliability of any biases recorded and standard error bars are shown on the graphs.

The bars labelled ‘Shift θ ’ and ‘bottom/top’ show results for when θ was shifted without the viewpoint being rotated, while the ‘front/back’ bars represent the situation when the scene was rotated away. We can see that performance was identical in both cases but that the hit-angle clearly had an effect. When θ was distorted upwards after the struck ball was hit from the bottom or downwards when hit from above (Figure 2(c), top), acceptance thresholds were significantly higher than in the opposite case ($p < 0.005$ and $p < 0.05$ respectively). When the viewpoint was rotated, the effects were almost identical i.e., when the ball was hit from the front and distorted away from the viewer or hit from behind and distorted towards, acceptance thresholds were significantly higher in the former case ($p < 0.05$) and marginally significant in the latter case ($p=0.06$). From the bars labelled ‘Shift β ’ we can see that significant effects ($p < 0.0005$ and $p < 0.005$) also occurred when β alone for the struck ball was distorted in the same way (Figure 2(d), top). The factor of hit-angle had no effect on the perception of distortions of the striking ball’s trajectory (Figure 2(e)), as seen from the bars labelled ‘Shift α ’, but clearly the acceptance thresholds for its distortions are far higher than for the struck ball ($p < 0.0001$). The factor of hit-angle also had no effect on the perception of expansion or contraction distortions, as shown by the grey bar labelled ‘Expand/Contract’, but the acceptance threshold was significantly higher for expansion of θ than contraction ($p < 0.05$).

The higher thresholds recorded showed that observers were almost 100% more accepting of angular distortions in our study than in Kaiser and Proffitt’s. Perhaps this is due to the fact that the objects were three-dimensional, thus introducing more uncertainty. The objects were also much bigger (4 cm diameter rather than 1cm) and the display was larger. The results were otherwise consistent. K&P only hit the stationary ball from above and found a higher acceptance for both shifting θ and expanding β downwards under this condition. They attempted to explain this as a bias towards gravity, even though their observers were told to discount it. However, as we found an equal and opposite effect when the stationary ball was hit from below i.e., a preference towards shifting θ and expanding β upwards, we can therefore conclude that this is not due to any bias towards the gravitational direction. We also tested for a rotated viewpoint and found an identical pair of biases depending on whether the struck ball was hit from the front or the back. The preferred distortion therefore corresponds to an expansion of the angle β rather than a bias towards any fixed direction, such as gravity. As in our study, Kaiser and Proffitt found no preferences for distortion direction of the striking ball’s trajectory and also recorded higher acceptance thresholds for α . However, our results demonstrate a strong preference for expansion of θ over contraction, a bias not revealed in the previous study. The biases we recorded were all statistically reliable.

From these results, we draw the following conclusions: People preferred expansion of θ over contraction, equivalent to an expansion of both α and β ; They also preferred the shifting of the whole system away from the striking ball’s pre-collision trajectory, equivalent to an expansion of β and a contraction of α . However, they showed no preference for expansion or contraction of α , but a strong preference for expansion of β . Therefore, we conclude that the bias in all of the θ distortions must be due mainly to the struck ball’s post-collision trajectory.

3.2.2 Momentum distortions

In the second block of experiments, we first ran one experiment where we manipulated the post-collision velocity of both balls. Fourteen observers volunteered: 1 female and 13 male colleagues and students aged 20 to 30. Seven of the observers had completed the first block of experiments and seven were new³. The display

³To counterbalance the order of the blocks across observers and to test for learning effects, of which none were found

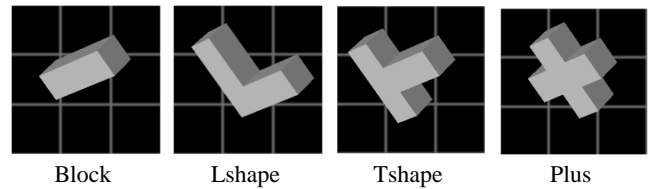


Figure 4: Objects for the experiment to evaluate distortions in angular velocity

and procedure were as described above, but only one hit-angle (25°) was used (as there is no evidence to suggest that this affects velocity discrimination) and only one initial velocity for the striking ball (12cms/sec) was tested. This gave rise to 4 randomly interleaved staircases.

The thresholds for linear velocity distortions are shown in the first bar of Figure 3(b). We can see that a speedup in the post-collision linear velocity of the balls was more acceptable to observers than a slow-down ($p = 0.05$).

3.2.3 Spatio-temporal distortions

The remaining three experiments of the second block of experiments were as follows:

1. *Early launch, Gap obscured* - We allowed the struck object to move before the collision actually occurred and we methodically increased or decreased the size of the gap between the objects at the time of this ‘early launch’. However, the viewpoint was rotated by almost 90° , thus occluding the gap for all but the most extremely distorted conditions. We tried two different approaches to determine the physics for these collisions, referred to as with or without proxies. In the ‘with proxies’ condition, we expanded the size of the objects as if they were touching and calculated the physics from these larger objects. In the ‘without proxies’ condition, the physics were calculated as if the objects themselves had actually touched i.e., by translating the struck ball further along its pre-collision trajectory until it touched the second ball and calculating the response based on these positions. This gave rise to 4 randomly interleaved staircases, ascending or descending for each of the with and without proxies conditions.
2. *Early launch, Gap visible* - This experiment was identical, except that the viewpoint was no longer rotated and any gap was therefore visible. Pilot experiments indicated that participants found that the grid in the background made it difficult for them to judge these collisions, so for this experiment the grid was removed.
3. *Causality* - We tested the causality effect by introducing a time delay at the moment of collision. We also hypothesized that the velocity of the striking ball would affect the perception of this delay, so we tested this at two different velocities as before, giving rise to 4 randomly interleaved staircases.

Figure 3(c) shows results for the spatio-temporal distortions. Unlike Kaiser and Proffitt, who recorded no tolerance whatsoever for their ‘early launch’ condition, we found that there are positive acceptance thresholds for some gaps. This is consistent with our previous studies [O’Sullivan and Dingliana 2001]. As we hypothesized, viewpoint had a highly significant effect, with acceptance thresholds being far higher for the ‘Gap obscured’ condition than for ‘Gap visible’ ($p < 0.0001$). The effect of using proxies for collision response was not significant in either condition. The effect of velocity on the perception of a time delay was also statistically insignificant, even though 14 out of 15 observers had a higher acceptance threshold for the slower condition than the faster. Perhaps

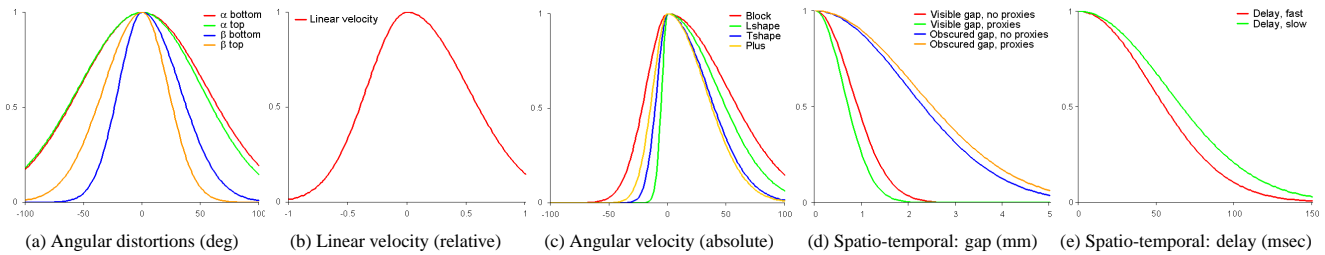


Figure 5: Gaussian probability functions for effects observed

a future study with a greater number of participants may lead to more reliable results. We also found a higher acceptance of delay than K&P, with 60msec delays being accepted on average.

3.2.4 Angular velocity distortions

Fourteen new observers were paid to participate in the final block of experiments: 10 male and 4 female undergraduate students, ranging in age from 19 to 28. Display was similar, except that the stationary object was one of 4 different models of varying complexity (Figure 4) and the scene was slightly rotated (10°) around the X and Y axes. The striking ball hit the stationary object at a point that imparted a spin and angular velocity was proportionally added or subtracted. There were 16 randomly interleaved staircases, one for each model type.

Thresholds for angular velocity distortions are shown in Figure 3(b)⁴. From the grey bars we can see that, as the complexity of the objects increased (and hence the number of edges that passed by at each rotation), the observers became less accepting of increases in angular velocity and more accepting of decreases. The exception was the block object, which had the highest acceptance thresholds for both distortions. The effect of the slow-down was significant across all four objects ($p < 0.05$, dof=3), while the effect of the speedup was not reliable (there was a high variation among subjects, who complained more about the difficulty of this task than any other). However, when the simplest and the most complex objects were compared, the effect of speed-up was marginally significant ($p < 0.1$). Further studies would be necessary to determine the reliability of this factor and to broaden the scope, but the results are consistent with previous studies [Barraza and Grzywacz 2002]. For now we will assume that object complexity, measured by the number of potential edge transitions, does have an effect on angular velocity distortion discrimination.

We can also see that there is a strong preference for an increase in angular velocity over a decrease ($p < .0001$), as with linear velocity. Acceptance thresholds for angular velocity distortions are higher than for linear velocity anomalies, probably because the event depicted involves the transformation from linear momentum to angular momentum and, as Proffitt and Gildea have shown, people do not have very adept appreciations of extended body systems. These events were also somewhat different than the simpler ones in the other experiments. When the two-sphere collisions are distorted, there is no set of relative masses that would make the event correct. Given the angle of incidence and the post-collision velocities, no possible relative weights would preserve the pre-collision momentum. For the sphere colliding with the extended body, the ‘anomalous’ post collision velocity would be correct were the object of a different density. So, although we told observers to assume the object has the same density as the sphere, this is a cognitive constraint

⁴Both the relative and absolute angular velocities are shown, as the relative results can be misleading. Unlike in the linear velocity case, for angular velocity the step-size was not identical for each of the four complex objects because the objects were of un-equal mass (but of the same material).

that the observer needs to impose on the situation and they therefore found this task to be the most difficult one⁵.

3.3 Proposed Metrics

Now that we have established some acceptance thresholds for the perception of certain types of dynamic anomalies and identified some factors that affect people’s perception of those anomalies, we need to extrapolate from these results a means of estimating the quality of a simulation that displays dynamic events. As in Cheney and Forsyth [2000], we will use a similar concept of the ‘probability’ of a simulation. A high probability value means that the animation is visually plausible, a value of 1 being fully plausible with respect to a given referent for reality. We know the X_{50} estimates for each type of anomaly i.e., the magnitude of distortion at which a person will see that anomaly 50% of the time. We estimate the probabilities of all other magnitudes by fitting a Gaussian probability function to this data. We achieve this as follows. If θ is the known threshold value, we want to determine the standard deviation δ for the probability function:

$$p(\theta) = \frac{1}{\sqrt{e^{(\frac{\theta}{\delta})^2}}} = \frac{1}{2} \Rightarrow \delta = \sqrt{\frac{\theta^2}{\ln(4)}} \quad (1)$$

As we have seen, these distributions are not necessarily symmetric, so we use two different standard deviations for each pair of opposite anomalies e.g., expansion vs. contraction. The probability functions for all anomalies tested are shown in Figure 5. Based on the data that we derive from our applications, we need to decide which of these functions can be used and also how they should be combined to derive some form of quality measure. We view the probability, or visual fidelity measurement, of an animation to consist of three components:

$$P(A) = f(P_{angular}(A), P_{momentum}(A), P_{spatiotemporal}(A)) \quad (2)$$

Let us consider each of these contributing probabilities separately. We have seen that the biases for angular distortions are probably due to distortions in the struck ball’s trajectory, but we need to generalize this in some way for situations where both colliding objects are in motion pre-collision, or when there is only one ball bouncing off a static object, such as a floor or walls. It is important to note that these are only assumptions, to guide the development of a metric, and we do not state them as facts. Further psychophysical experiments would be advisable in order to validate their reliability. We make the following major assumption that allows us to extrapolate: the distortions in the struck ball are more noticeable because the viewer is attending its trajectory and not that of the striking ball. Hence, the striking ball simply becomes another distracting item,

⁵Although there is some evidence that people tend to assume homogeneity of density, at least when looking at a single object [Carlton and Shepard 1990], there is no evidence to show that this applies across different objects in the scene.

albeit a colliding one. All other visible but unattended colliding objects are treated in the same way as the striking ball i.e., they exhibit no bias in terms of the direction of angular distortion. We will call the probability function to be applied to the single struck/attended ball P_β and the striking/unattended balls, P_α . Hence, the probability associated with the post-collision trajectory of any object x_i involved in a collision is:

$$P_{angular}(x_i) = P_L(x_i)P_\beta(\theta_i) + P_{NL}(x_i)P_\alpha(\theta_i) \quad (3)$$

where θ_i is the angle of distortion and $P_L(x_i)$ and $P_{NL}(x_i)$ are the probabilities that the viewer is looking, or not looking, at object x_i . If there is only one object bouncing around, then $P_L(x_i)$ is 1 and $P_{NL}(x_i)$ is 0. If there are two objects colliding, then if we can determine by some means which ball is being attended e.g., if it was stationary before collision or if eye-tracking is used, then the probability for the attended ball becomes $P_\beta(\theta_i)$ and for the unattended ball is $P_\alpha(\theta_i)$. If both balls were moving and we don't know which one was attended, $P_L(x_i)$ and $P_{NL}(x_i)$ are each 0.5, while in the presence of D distractors, they become $1/D$ and $1 - 1/D$ respectively. We now suggest a formulation for the angular component of the visual plausibility metric of an animation A to be a function of the collision normals for all objects involved in collisions:

$$P_{angular}(A) = f_A(P_{angular}(x_i)) \forall i \quad (4)$$

The momentum distortion is described similarly:

$$P_{momentum}(A) = f_M(P_{LV}(x_i), P_{AV}(x_i)) \forall i \quad (5)$$

where the linear and angular velocity contributions P_{LV} and P_{AV} are shown in Figure 5 (P_{AV} is selected based on the number of potential edge transitions). Finally, the gap factor relates to each collision pair c_j , while the delay factor affects all collisions equally, so we describe the spatio-temporal component to be:

$$P_{spatiotemporal}(c_j) = f_{ST}(P_{gap}(c_j), P_{delay}) \forall j \quad (6)$$

We are not prescriptive as to how these functions are defined. This is controlled by the concept of ‘fitness for purpose’ described above (perhaps some anomalies are more or less acceptable than others in specific applications). Also, our simple experiments do not yet allow us to claim with authority how these perceptual factors interact and thus to derive relative weights. Instead, just as Daly does not output a single number to describe the similarity of two images but produces a 2-dimensional map of the visible differences, we will also examine these individual components in each of our case studies. Hopefully, this will guide others in choosing the metrics they can use, or indeed deriving their own if these are not sufficient for their purposes.

4 Case Studies

In order to demonstrate the applicability of the metric, we present two case studies - simulation levels of detail and constrained dynamics.

4.1 Simulation levels of detail

We use a system that supports simulation levels of detail by using collision proxies in an adaptive mechanism for collision handling. A hierarchy of spheres is used to provide successively closer approximations to a rigid body for collision detection (Figure 6). These increasing levels of detail are used to adaptively improve the quality of collision response [Dingliana and O’Sullivan 2000]. Exploiting such a system, we show how the visual fidelity functions proposed in Section 3.3 could be used for evaluation purposes.

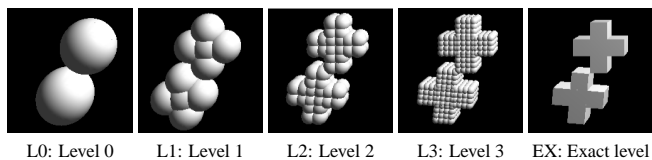


Figure 6: The hierarchical collision proxies used for the Simulation Levels of Detail case study.

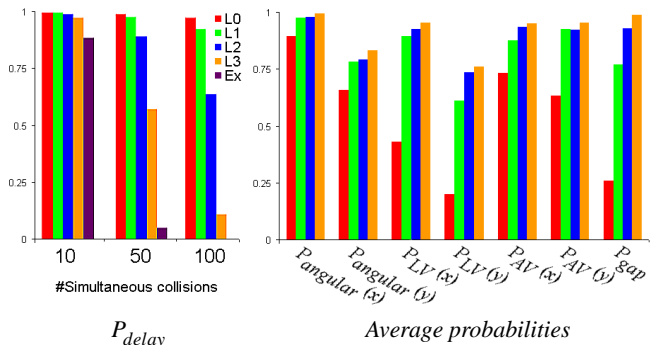


Figure 7: Probabilities for the Simulation Levels of Detail case study - L0 to EX are shown above

Pairs of rigid bodies (those used in Block 3), 100 for each level, were given random starting states and set to collide in a frictionless environment without gravity. One body, which we will call body x , was initially moving while the other body, y , was initially stationary. The *state* of a rigid body here refers to its linear and angular velocity as well as its position and orientation. For each simulation, we recorded both the pre-collision states of the two objects and also the post-collision states as computed by the approximate collision handler. Note that different resolutions of collision handling lead to different post-collision states in the colliding bodies and the system is designed to adaptively reduce this form of *unavoidable distortion* as more time becomes available for collision processing. After the simulation was completed, we compared the collision data with a *referent* i.e., a collision handler that represents the exact solution. To do this we needed to calculate post-collision states of the objects based on the pre-collision states recorded during the simulation as accurately as possible.

Unfortunately, approximate collision detection generally leads to objects being treated as colliding even though a separation distance (or gap) remains between the two objects. The approximate collision handler is able (and indeed required) to calculate a response from such an approximate contact. However, in order for the exact collision handler to compute a response, it is required that we provide an equivalent ‘gapless’ state that the reference collision handler can deal with. We make the assumption that an equivalent gapless state is one in which the two bodies have been translated towards each other in the direction of least separation. Then, based on this hypothetical *translated state* it is possible to calculate an accurate response using the exact collision handler. It is debatable whether the accurate response from the translated state should necessarily be the basis for comparison. The other possible alternative would be to obtain the gapless state by continuing the simulation until actual contact is achieved and to use this instead as a basis for comparison. We justify our use of the former case by observing that the continuation of objects with angular velocity further along their accurate paths would give rise to a pair of relative orientations that could be completely different from when the approximate response was computed. It is unlikely that people mentally rotate objects forwards in time in this way (although further tests would be needed to confirm this).

The probability functions were applied to the data recorded and are shown in Figure 7. We recorded the time taken for every collision at all levels, including the exact level. These results were used to estimate the time needed to perform 10, 50 and 100 simultaneous collisions at each of the levels, which was in turn used to calculate P_{delay} using the average of the functions shown in Figure 5(e). We also calculated the average probabilities for the gap (P_{gap}) in each collision, along with the angular distortions ($P_{angular}$) and linear and angular velocity distortions (P_{LV} , P_{AV}) of the struck ball y and the striking ball x . These average probabilities give a good indication of the increasing visual fidelity of the animation with higher levels of detail, although the delay also increases due to the increasing computational load. By applying these functions to data recorded for similar tests with alternative hierarchical collision proxy models, as discussed in Section 2.2, this would allow for a comparison of different techniques in terms of both efficiency and visual fidelity.

4.2 Constrained dynamics

We developed a simple physically based system that simulates two balls both influenced by gravity starting at an initial position in the air. They are then sent towards each other so that they collide before hitting the plane below them (Figure 8). We direct the colliding balls towards desired spots on the plane, by changing only the linear component of the resultant momenta. This is enough for us to gradually guide the balls in the desired direction so that they finally end up at, or close to, the desired positions. After the initial two ball collision, each time a ball hits and bounces off the plane, it is further directed towards the desired spot using constraints [Barzel and Barr 1988], which are applied only for an instantaneous moment of time, enough to suggest to the ball that it should be going in a particular direction rather than the direction that is warranted by the proper response of the collision. The velocity magnitude of the ball is also affected, causing a speed-up or slow-down of the ball after it strikes the plane. The balls will often not land exactly on the spot but close to it and we stop the simulation when either ball reaches a state classified as rest to avoid them rolling forwards and backwards trying to reach the end-point.

For each collision we record the original angles between the incoming and outgoing linear momentum, the new augmented angles due to the applied changes of the new outgoing momentum, and the original and newly augmented velocity magnitudes of each ball. To evaluate the angular distortion probability $P_{angular}$, we applied Equation 3 from Section 3.3. The probability of looking P_L or not looking P_{NL} at each ball is 0.5. The objects did not have any angular velocity distortions, so the momentum probability $P_{momentum}$ was calculated using the linear velocity probability P_{LV} alone and there were no spatio-temporal distortions. The results are shown in Table 1, in the form of probability averages, products (as used in Cheney and Forsyth [2000]) and minimums. The averages are similar, whereas the minimums and products would suggest that the animation in Scene 2 has a higher visual fidelity. The interpretation of this information is up to the developer or designer. It could be used, for example, to improve an algorithm to generate only animations with a lowest error bound, or to choose an aesthetically distorted simulation for a game or movie.

5 Conclusions and Future Work

In this paper we have identified the need for a metric to measure the visual fidelity of animations and proposed some new functions that could be used. This work is clearly just a first step towards such a metric, but we hope that it will form a basis for future investigations. As they stand, the proposed probability functions are only valid for measuring the visual fidelity of the simple events described. We cannot yet predict how these metrics will generalize to more complex animations, involving multiple objects undergoing

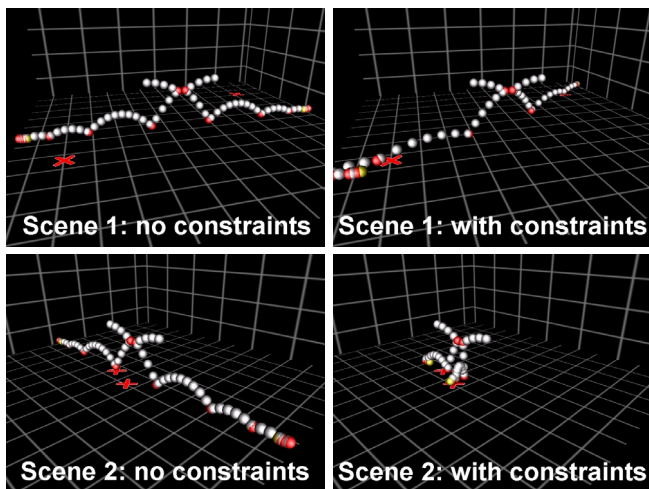


Figure 8: Constrained dynamics framework - red X marks show the constraints that ‘pull’ the ball towards that position.

	Scene 1		Scene 2	
	$P_{angular}$	$P_{momentum}$	$P_{angular}$	$P_{momentum}$
Product:	0.01	0.06	0.17	0.14
Average:	0.89	0.84	0.88	0.85
Minimum:	0.02	0.37	0.35	0.52

Table 1: Probabilities for the Constrained Dynamics case study.

more complex interactions with friction, gravity, resting and sliding contacts. Animation consists of many separately perceivable components and it is not yet safe to assume that these thresholds can be measured and applied independently. As stated in Section 3.3, we do not yet know how the factors tested interact with each other e.g., whether angular velocity will affect the perception of linear velocity, or with other factors such as attention and experience. However, as Proffitt and Gilden [1989] demonstrated, when the complexity of physical systems increases, people become less competent at noticing dynamic anomalies. It remains to be seen if this holds when further factors are added, such as sound, texture and behavior and we intend to investigate these issues further and encourage others to do the same. We believe that a probability-based approach to evaluation, as described in this paper, is ideal for modelling the interaction of these perceptual parameters.

Also relevant may be concerns that have been raised regarding the generalizability of automatic image fidelity metrics that depend on error sensitivity measurements. For example, Wang et al. [2002] cite an online report by the Video Quality Experts Group⁶, which reported that the performance of many complicated image metrics was no better than simple statistical predictors such as Peak Signal-to-Noise Ratio (PSNR). We feel far more optimistic, however, about the potential for generalizing from the results in this paper to a more generic animation metric. Image metrics, to date, do not use any saliency information about the scene being depicted and rely on predictions based on quite low-level responses in the human visual system (HVS). Wang et al. note that the HVS is highly adapted to extract structural information from the viewing field and recommend using a measurement of structural distortion to model image degradations instead of the more common error sensitivity based approach. When working with 3D animation systems, the evaluation system has access to this structural information, by definition, along with much more saliency information that can be used to determine attention and importance.

⁶Final report March 2002, available from <http://www.vqeg.org/>

6 Acknowledgements

We would like to acknowledge the support of the Higher Education Authority of Ireland. Thanks also to all who participated in our study and to the anonymous reviewers for their helpful comments on the presentation and content of this paper.

References

- BARRAZA, J. F., AND GRZYWACZ, N. M. 2002. Measurement of angular velocity in the perception of rotation. *Vision Research* 42, 2457–2462.
- BARZEL, R., AND BARR, A. H. 1988. A modelling system based on dynamic constraints. In *Computer Graphics (Proceedings of ACM SIGGRAPH 88)*, ACM, 179–188.
- BARZEL, R., HUGHES, J. F., AND WOOD, D. N. 1996. Plausible motion simulation for computer graphics animation. In *Computer Animation and Simulation* 96, 184–197.
- BROGAN, D. C., AND HODGINS, J. K. 2002. Simulation level of detail for multiagent control. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems*, 199–206.
- CARLSON, D., AND HODGINS, J. 1997. Simulation levels of detail for real-time animation. In *Proceedings Graphics Interface*, 1–8.
- CARLTON, E. H., AND SHEPARD, R. N. 1990. Psychologically simple motions as geodesic paths: I. asymmetric objects. *Journal of Mathematical Psychology* 34, 127–188.
- CHENNEY, S., AND FORSYTH, D. 1997. View-dependent culling of dynamic systems in virtual environments. In *Proceedings of the ACM Symposium on Interactive 3D Graphics*, 55–58.
- CHENNEY, S., AND FORSYTH, D. 2000. Sampling plausible solutions to multi-body constraint problems. In *Proceedings of ACM SIGGRAPH 2000*, ACM Press / ACM SIGGRAPH, 219–228.
- CHENNEY, S., ARIKAN, O., AND FORSYTH, D. 2001. Proxy simulations for efficient dynamics. In *Proceedings of Eurographics 2001, Short Presentations*.
- CLEMENT, J. 1982. Students' preconceptions in introductory mechanics. *American Journal of Physics* 50, 1, 66–71.
- CORNISWEET, T. 1962. The staircase method in psychophysics. *American Journal of Psychology* 75, 485–491.
- DALY, S. 1993. The visible differences predictor: an algorithm for the assessment of image fidelity. In *Digital images and human vision*, 179–206.
- DINGLIANA, J., AND O'SULLIVAN, C. 2000. Graceful degradation of collision handling in physically based animation. *Computer Graphics Forum (Eurographics 2000 Proceedings)* 19, 3, 239–247.
- EHMANN, S. A., AND LIN, M. C. 2001. Accurate and fast proximity queries between polyhedra using convex surface decomposition. *Computer Graphics Forum (Eurographics 2001 Proceedings)* 20, 3, C500–C510.
- FUNKHOUSER, T., AND SÉQUIN, C. . 1993. Adaptive display algorithm for interactive frame rates during visualization of complex virtual environments. In *Proceedings of ACM SIGGRAPH 1993*, ACM Press / ACM SIGGRAPH, 247–254.
- GILDEN, D., AND PROFFITT, D. 1989. Understanding collision dynamics. *Journal of Experimental Psychology: Human Perception and Performance* 15, 2, 372–383.
- GOTTSCHALK, S., LIN, M., AND MANOCHA, D. 1996. Obb-tree: A hierarchical structure for rapid interference detection. In *Proceedings of ACM SIGGRAPH 1996*, ACM Press / ACM SIGGRAPH, 171–180.
- GRAZIANO, M. S. A., ANDERSON, R. A., AND SNOWDEN, R. J. 1994. Tuning of mst neurons to spiral motions. *J. Neuroscience* 14, 1, 54–67.
- GROSS, D. 1999. Report from the fidelity implementation study group. In *Fall Simulation Interoperability Workshop Papers*.
- HODGINS, J., O'BRIEN, J., AND TUMBLIN, J. 1998. Perception of human motion with different geometric models. *IEEE Transactions on Visualization and Computer Graphics*. 4, 4, 307–316.
- HUBBARD, P. 1995. Collision detection for interactive graphics applications. *IEEE Transactions on Visualization and Computer Graphics*. 1, 3, 218–230.
- KAISER, M. K., AND PROFFITT, D. R. 1987. Observers' sensitivity to dynamic anomalies in collisions. *Perception and Psychophysics* 42, 3, 275–280.
- KLOSOWSKI, J., HELD, M., MITCHELL, J., SOWIZRAL, H., AND ZIKAN, K. 1998. Efficient collision detection using bounding volume hierarchies of k-dops. *IEEE Transactions on Visualization and Computer Graphics* 4, 1, 21–36.
- LEVITT, H. 1971. Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America* 49, 467–477.
- LONGRIDGE, T., BRKI-COHEN, J., GO, T. H., AND KENDRA, A. J. 2001. Simulator fidelity considerations for training and evaluation of today's airline pilots. In *Proceedings of the 11th International Symposium on Aviation Psychology*, The Ohio State University Press.
- LUEBKE, D., AND HALLEN, B. 2001. Perceptually driven simplification for interactive rendering. In *Rendering Techniques*, Springer-Verlag, London, S. Gortler and K. Myszkowski, Eds., 223–234.
- MICHOTTE, A. 1963. *The Perception of Causality*. Basic Books, New York.
- MIRTICH, B. 2000. Timewarp rigid body simulation. In *Proceedings of ACM SIGGRAPH 2000*, ACM Press / ACM SIGGRAPH, 193–200.
- MYSZKOWSKI, K., TAWARA, T., AKAMINE, H., AND SEIDEL, H.-P. 2001. Perception-guided global illumination solution for animation rendering. In *Proceedings of ACM SIGGRAPH 2001*, ACM Press / ACM SIGGRAPH, 221–230.
- O'BRIEN, D., FISHER, S., AND LIN, M. C. 2001. Automatic simplification of particle system dynamics. In *Proceedings of Computer Animation, 2001*.
- O'SULLIVAN, C., AND DINGLIANA, J. 2001. Collisions and perception. *ACM Transactions on Graphics (TOG)* 20, 3, 151–168.
- POPOVIC, J., SEITZ, S. M., ERDMANN, M., POPOVIC, Z., AND WITKIN, A. 2000. Interactive manipulation of rigid body simulations. In *Proceedings of ACM SIGGRAPH 2000*, ACM Press / ACM SIGGRAPH, 209–217.
- PROFFITT, D., AND GILDEN, D. 1989. Understanding natural dynamics. *Journal of Experimental Psychology: Human Perception and Performance* 15, 2, 384–393.
- RAMASUBRAMANIAN, M., PATTANAIK, S. N., AND GREENBERG, D. P. 1999. A perceptually based physical error metric for realistic image synthesis. In *Proceedings of ACM SIGGRAPH 1999*, ACM Press / ACM SIGGRAPH, 73–82.
- ROZA, M., VOOGD, J., JENSE, H., AND VAN GOOL, P. 1999. Fidelity requirements specification: A process oriented view. In *Fall Simulation Interoperability Workshop*.
- ROZA, M., VAN GOOL, P., AND VOOGD, J. 2000. Fidelity considerations for civil aviation distributed simulations. In *Proc. AIAA Modeling and Simulation Technologies Conference*.
- WANG, Z., BOVIK, A., AND LU, L. 2002. Why is image quality assessment so difficult? In *Proceedings of the IEEE International Conference on Acoustics, Speech, & Signal Processing*, vol. 4, 3313–3316.
- WATSON, B., FRIEDMAN, A., AND MCGAFFEY, A. 2001. Measuring and predicting visual fidelity. In *Proceedings of ACM SIGGRAPH 2001*, ACM Press / ACM SIGGRAPH, 213–220.
- YEE, H., PATTANAIK, S., AND GREENBERG, D. P. 2001. Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. *ACM Transactions on Graphics* 19, 2, 39–65.