Flexible Rendering of 3D Graphics Under Varying Resources: Issues and Directions

Eric Horvitz and Jed Lengyel

Microsoft Research
Redmond, Washington 98052
{horvitz, jedl}@microsoft.com

Abstract

The Talisman project centers on the creation of a flexible multimedia architecture that can deliver high-performance 3D graphics on personal computing platforms. The architecture hinges on the use of flexible graphics strategies that exploit the coherence of objects adjacent frames. We describe the Talisman approach, focusing on flexible strategies that provide for tradeoffs between the quality and computational costs of rendering, and review key challenges and opportunities for enhancing the selective allocation of resources in the system.

1 Introduction

The Talisman project at Microsoft [Torborg and Kajiya, 1996] is focused on the creation of a graphics rendering architecture with the ability to deliver high-quality interactive graphics on inexpensive personal computing platforms. The architecture harnesses flexible procedures for delivering graphics, audio, and video under varying resource requirements and limitations. We will focus on the control of tradeoffs in the graphics component of Talisman. Analogous methods for introducing flexibility have application for the control of tradeoffs in audio and video processing within the Talisman system.

Talisman decomposes the problem of rendering scenes into a set of subproblems centering on the rendering of animated subcomponents of objects called *sprites*. The system attempts to minimize computational needs by making decisions about the best way to handle tradeoffs with rendering

Presented at *Symposium on Flexible Computation*, M.I.T., Cambridge MA,
Report FS-96-06, AAAI: Menlo Park, CA, November 1996, pages 81-88.

individual sprites. Sprites are generated as individual image layers and are composited together to generate the final output. During the construction of each frame, the image layers are traversed and decisions are made about the display of each of the layers. Thus, individual sprites can be processed separately and can be subjected to a variety of flexible approximations.

A primary motivation behind the design of Talisman is to take advantage of temporospatial coherence in adjacent frames. Such coherence arises from the fundamental stabilities of the structural composition of objects, the typical continuity of trajectories of objects in motion, and smoothness in the dynamics of viewpoint. Talisman attempts to minimize the computational requirements of rendering by employing inexpensive transformations on previously rendered sprites rather than re-rendering sprites where possible. By analyzing and harnessing interframe coherence, Talisman moves graphics from an endeavor focused purely on image synthesis to a combination of image synthesis and image processing.

The Talisman approach to decomposing the graphics problem into a fine-grained set of sprite-rendering subproblems provides opportunities for the variable allocation of resources to the solution of subproblems that promise to contribute the most to the overall quality of the rendered frames. We will review issues with the selective control of flexible algorithms to solve these subproblems. In particular, we will describe the regulation of tradeoffs between the computation required to solve graphics subproblems and the quality of the final images.

Research on the selective allocation of resources in Talisman comes in the context of related work on the control of tradeoffs in graphics systems. Our concerns and goals are closely related to the work Funkhouser and Sequin, who have explored adaptive strategies for selectively controlling the level of detail of objects in the interactive visualization of architectural models, taking a similar cost–benefit perspective on graphics tradeoffs under resource constraints [Funkhouser and Sequin, 1993].

We will describe fundamental problems and opportunities with the control of multiple dimensions of degradation in a sprite-oriented system. After reviewing several flexible-rendering strategies, we will dwell on classes of information required for the effective control of resources allocation in the system. Then we will focus on the challenges of characterizing the relationships between the computational costs associated with interrelated phases of computational analysis and the perceptual quality of alternative solutions.

2 Introducing Flexibility into Graphics Rendering

Traditional approaches to rendering graphics employ a graphics processing pipeline that renders objects on a frame-by-frame basis. The graphics pipeline begins with a specification of the geometry of model objects and the environment in a scene. Objects are described by vertices that describe polygons composing the surfaces of models, by normals to surfaces of the objects, and with texture coordinates, colors, and texture maps. The environment is described by geometrical transformations based on the models, the viewpoint, and lighting and shading parameters. The rendering systems integrate information about model and environment to ultimately rasterize and generate video through the production of pixels.

In contrast with Talisman, traditional approaches to rendering adhere to predetermined, static levels of quality of rendered objects. Performance in these systems is measured in terms of the number of polygons processed each second. Thus, slower machines are linked intrinsically to longer computation times for rendering the constant-quality images. By nature, these inflexible systems lead to an inescapable slowing of frame rates as computational resources are reduced or the complexity of images increases.

There are a variety of ways to introduce flexibility into the graphics pipeline. Talisman decomposes images into sprites and uses a set of flexible approximation procedures that allow the system to trade off the quality and cost of rendered objects. Strategies for selectively controlling these tradeoffs can endow the system with the ability to react robustly to variations in the overall resource requirements as scenes change in complexity or as available resources are diminished because of the changing pull on resources by other applications (including other multimedia tasks) running within a personal computing environment. The coupling of flexible strategies with effective regulation of tradeoffs can also minimize graphics degradation as content is rendered on platforms with diminished computational power.

3 Flexible Strategies and Dimensions of Degradation

Flexible computation strategies are procedures that allow a graceful tradeoff to be made between the quality of results and allocations of costly resources such as time and memory. Systems employing flexible strategies gain the ability to dynamically adapt the quality of their response to changes in requirements for precision, and to uncertainty or variation in the cost and availability of computational commodities. Desiderata of flexible computation include bounded discontinuity on the value of partial results generated as the allocation of resources decreases, and convergence on an ideal result with sufficient resources [Horvitz, 1987]. There have been a variety of studies on the use of flexible strategies in computational problem solving.

Flexible graphics rendering methods allow Talisman system to trade off the fidelity of components of images for computational savings. Each approximation induces a different type of degradation along a specific dimension of loss in fidelity as the quantity of allocated computation is reduced.

Flexible rendering methods include approximations that allow for the graceful reduction of model complexity, spatial resolution, shading complexity, and temporal resolution. For each of the degradation dimensions, there is a corresponding fiducial which measures or estimates the distance from a gold-standard or most accurate rendering along that dimension of quality.

Simplification methods based on the control of model complexity work at the level of the fundamental geometries that represent objects in a scene. The dynamic selection of simpler models from a spectrum of models to enhance rendering performance (referred to as *detail elision*) has been employed in a number of graphics systems. Model-simplification methods include the storage of multiple models, the dynamic simplification of the mesh of polygons that specify a model, and the choice of sampling for surfaces specified by parametric models. As the geometry used to describe a model is simplified, details of its structure are lost and artifacts appear in the form of errors in the overall shape or surface characteristics of objects.

An example of a flexible model-simplification strategy is the *progressive mesh* developed by Hoppe [Hoppe, 1996]. With this approach, edges of polygons composing a mesh are selected for collapse into single vertices so as to minimize a measure of energy defined in terms of multiple factors including the total number of vertices and the squared distance of points in the simplified models from an ideal mesh. The edge-collapse operation is reversible with tractable vertex-splitting operations, allowing for rapid traversal of a spectrum of meshes. Figure 1 displays a fully detailed model and a point on the spectrum of increasingly simplified models developed by the progressive mesh method. Figure 2 shows the relationship between the simplification of models with the progressive mesh method as measured by the resulting number of vertices, and a measure of model fidelity, as captured by the squared distance of points in the simplified model from the initial fully detailed mesh.

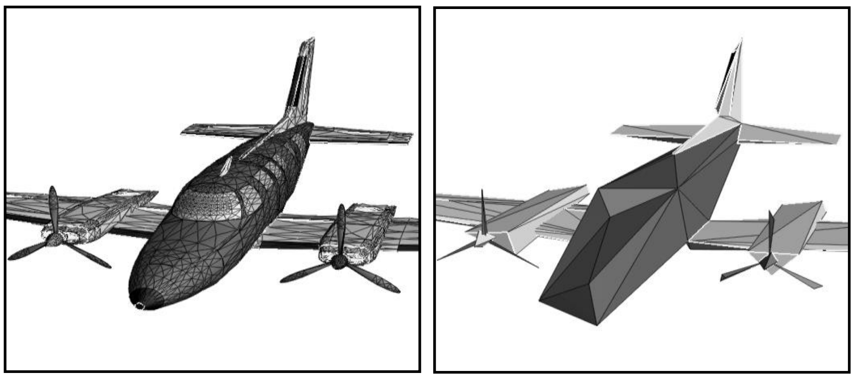

Figure 1: Flexible strategy for reducing the complexity of models. Methods for reducing the number of vertices used to describe a geometric model can be used to generate a spectrum of models from an ideal model to progressively simpler models. The fully detailed model (left) is described by 6,795 vertices. The simplified model (right) is described by 97 vertices.

Spatial resolution methods center on the control of the number of samples dedicated to a sprite and by the image compression factor. As indicated in Figure 3, we can selectively reduce the number of pixels devoted to a sprite. As the resolution is diminished, objects lose detail and ultimately become granular and fuzzy.

The shading complexity is determined by the type of texture filtering, texture level-of-detail, the number of lights used to illuminate a scene, and the use of shading effects such as reflections or shadows. Diminishing shading complexity introduces granularity and artifacts in the subtleties of lighting and reflection. The magnitude of error can be captured by photometric estimates.

Temporal resolution methods apply computationally inexpensive transformations to update sprites generated for previous frames so as to approximate the general motion of an object or change of viewpoint in three dimensions—instead of undertaking more expensive re-rendering of the sprites from 3D models. The manipulation of temporal resolution is central in Talisman. The system continues to attempt to adapt previously rendered sprites to new frames, bypassing the more expensive task of re-rendering the sprites, but potentially inducing spatial and temporal artifacts. In the general case of rendering multiple frames of an object in motion, or of changing viewpoint or lighting, the reuse of sprites may introduce errors because of the limited ability of translation, rotation, and warping with an affine transformation to simulate general three-dimensional changes. Several groups have explored the similar use of inexpensive warping transformations on

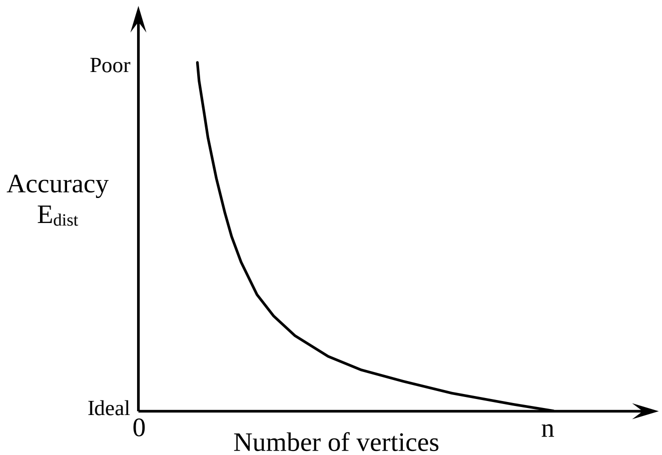

Figure 2: Tradeoff between size and accuracy. The graph displays how the relationship between a measure of model accuracy based on an energy quantity (E_{dist}), defined as the total squared distance of the points in a reduced model from the fully detailed mesh describing the ideal model, varies with the size of the model produced by the progressive mesh approximation procedure.

subcomponents of images to reduce real-time computation [Regan and Pose, 1994], [Shade96 et al., 1996], [Chen and Williams, 1993], [McMillan and Bishop, 1995].

In Talisman, sprites rendered in an earlier frame will continue to be transformed and reused in new frames until an estimate of error exceeds a tolerated error level. A small number of positions in the image, called characteristic points, are used to monitor and estimate the error as a function of the distances between the reused sprite and the actual sprite in the new frame. As the level of tolerated error is increased, the number of sprites that need to be re-rendered in each frame typically drops. However, as the error threshold is increased, geometric and temporal artifacts become more salient as sprites that have been warped for several frames are replaced with a re-rendered sprite after the larger error thresholds have been exceeded. By increasing the tolerated error, we save on resources by minimizing re-rendering, but may introduce noticeable artifacts in object distortion, problems with visibility of overlapping sprites, and discontinuities.

Figure 4 highlights the savings that can be achieved with sprite reuse. The left panel of the figure displays a frame drawn from a sequence of frames produced by a Talisman simulation that depicts two spacecrafts traveling over mountainous terrain. The right panel shows a set of boxes that bound the individual sprites in the scene. The sprites that have been re-rendered for this frame are highlighted with solid lines, while the larger number of sprites that have been reused through inexpensive transformations are bounded with rectangles composed of broken lines.

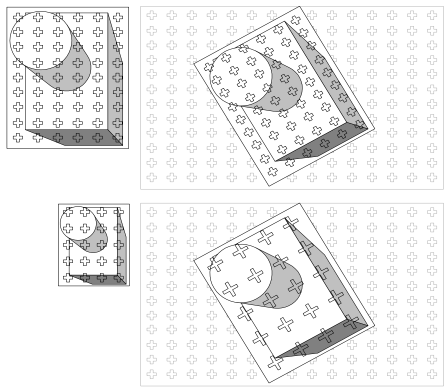

Figure 3: Reducing the spatial resolution. Objects can be sampled at a variety of spatial resolutions before being composited into final scenes, allowing for a tradeoff between the clarity of an object and the computation required to render the object.

4 Monitoring and Control of Graphics Computation

Given a set of flexible strategies that allow a system to manipulate tradeoffs between the computational requirements of rendering and the fidelity of the scenes, we seek policies that will optimally allocate resources. As we increase the degree of approximation, we must consider the *gains* in computational savings and the *losses* in the perceived quality of images. In addition to considering single dimensions of degradation, we must also characterize the computational and perceptual implications of using multiple approximations and engaging in parallel several dimensions of potentially interrelated degradations.

4.1 Characterizing Computational Needs

Controlling the allocation of resources relies critically on estimates of the cascade of computational costs as sprites are rendered from models. The cost of specific combinations of approximations is a function of interdependent processes at different locations in the graphics pipeline. For example, simplifying the geometry of an object early in the graphics pipeline can have dramatic effects on resource requirements because of the magnification of savings at different phases of graphics processing. Simplifying models will typically effect the computation required for shading, reflection, and re-rendering.

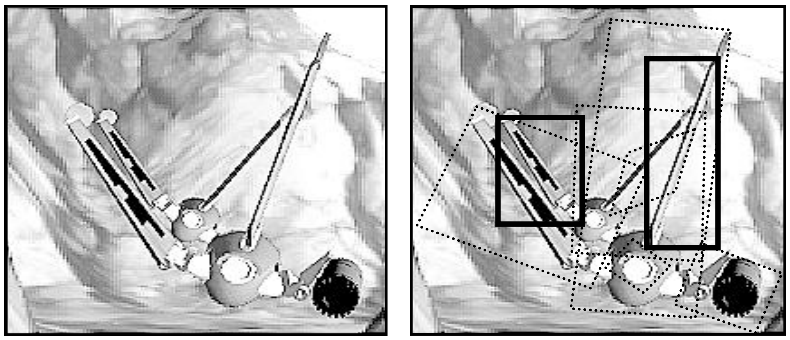

Figure 4: Minimizing computation via sprite reuse. Talisman takes advantage of spatial and temporal coherence by transforming previously rendered sprites instead of re-rendering the sprites. Rectangles composed of dotted lines bound sprites warped to fit a new frame; rectangles of solid lines bound re-rendered sprites.

Interactions among different approximations can be difficult to characterize. For example, degrading a model via a model-simplification strategy may dramatically decrease the number sprites required to represent the object. However, the frequency with which the sprites need to be re-rendered at the same level of error tolerance may increase.

Previous research on solutions to the problem of ideally partitioning resource to interdependent phases of computation has relevance to the problem of partitioning resources to different processes in the graphics pipeline. Efforts have explored the case of deterministic and uncertain relationships in the efficacy of interdependent phases of analysis [Horvitz and Breese, 1990], [Breese and Horvitz, 1990], [Dean and Wellman, 1991], [Zilberstein, 1993].

Although models of the interdependence of phases of computation for optimizing the control of rendering may be crucial, the details of relationships among strategies may evade modeling a priori because of complex context-dependent relationships. Thus, we may be forced to rely on real-time monitoring to probe *scene-specific* computational resource relationships with real-time monitoring. Talisman has a set of monitors for gathering information about the cost of computation.

4.2 Modeling Perceptual Quality

The complement to characterizing the costs of computation is the development of models capturing the costs of approximations on the *perceived quality* along single and multiple dimensions of graphics degradation. The overall perceptual quality of rendered images or the specific cognitive costs associated with various types and magnitudes of artifacts may be complex functions of the specific content of a scene. For example, the system may enhance the complexity of the surface of objects by passing a more complex geometric model down the graphics pipeline but may be forced to increase the tolerated error between affine transformations and actual sprites to maintain the frame rate. In this case, we would be forced to grapple with the perceptual tradeoff between the overall richness of a model and the discontinuities associated with more frequent and significant jumps when sprites are re-rendered. Elucidating general relationships between the nature and degree of approximations and the quality of final scenes is critical in the effective regulation of a flexible rendering system.

We may have easy access to information of the form portrayed in Figure 2, detailing the relationships between simple summaries of errors and alternate simplifications generated by flexible strategies. However, to understand the actual perceptual costs and benefits, we need to make the additional link to computational resources and perceptual accuracy, and to extend our understanding to multiple combinations of degradation. This goal is captured by Figure 5, which highlights the missing links between measures of accuracy and estimates of computational load, and between measures of accuracy and the perceived quality of images.

The overall aim of capturing the influence of multiple factors on the perceptual quality of images is the development of a rich, multiattribute utility model that represents the perceptual cost associated with the degradation of images from a gold-standard image. The perceptual cost associated with an frame rendered with a variety of degradations, relative to a fully detailed image created when there is sufficient resources to render a fully detailed frame, is typically a function of attributes that characterize the status of sprites within and across frames. Because Talisman relies on spatial and temporal approximations, its utility model should operate on attributes describing the degradation of multiple sprites within single frames as well as attributes from frames at different times (*e.g.*, attributes capturing motion artifacts and temporal discontinuities), as represented by the following cost function,

$$\text{Cost} = C[a_1(t_o),...a_{n-1}(t_o),\ a_n\ [\ a_i\ (t_o),\ a_j(t_{-1})]] \qquad (1)$$

where $a_i\ (t_o)$ represents attributes within a current frame and $a_i(t_{-1})$ represents attributes within the prior frame.

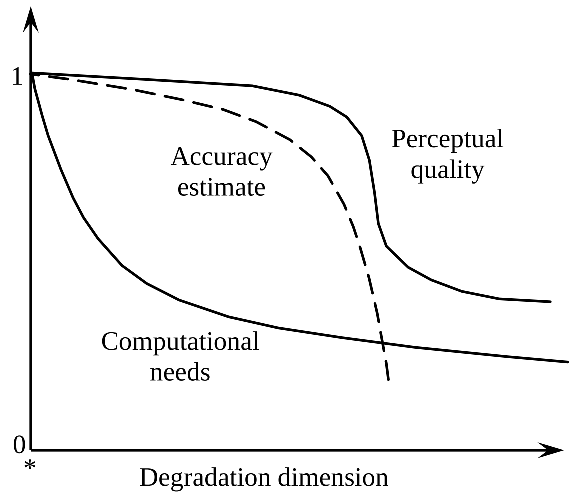

Figure 5: In pursuit of perceptual costs and computational savings. We seek to move beyond information about the relationships between estimates of accuracy and measures of simplification, to gain a better understanding of the overall perceptual losses and computational savings as graphical content is degraded along one or more dimensions.

Related work on the control of computation has demonstrated how multiple dimensions of value in a computational result can be characterized with multiattribute utility models. Such models have been employed to monitor and control the refinement of multiple attributes of value in the results generated by flexible algorithms for such tasks as sorting and decision making [Horvitz, 1988], [Horvitz, 1990].

4.3 Greedy and Global Optimization Strategies

The ultimate goal of characterizing the tradeoffs between computation and perceptual quality of images is to collect knowledge to support decisions to selectively degrade aspects of image associated with the greatest computational gain and minimal perceived loss in image quality.

Achieving this goal requires us to understand both the interdependencies among approximations and the human perceptual sensitivities.

Several forms of control are feasible, from simple, greedy approaches to more complex optimizations that consider a large number of configurations of allocations. The overall architecture and compositing methodology in Talisman creates an amenable environment for experimenting with a variety of greedy algorithms for allocating resources to re-rendering of individual sprites. We can allocate small quantities based on the costs and benefits associated with the current re-rendering. Greedy prioritization methods can bypass costly dynamic real-time optimization, allocating resources in the best way for a set of small amount of resources.

A default method for allocating resources in initial versions of Talisman is a greedy cost–benefit scheduler. For each sprite, the system sorts the sprites according to a function of the costs and benefits associated with minimizing the error as estimated by the flexible graphics strategies. The cost of a sprite is a weighted measure of the error and the importance of sprites as assigned by the authors of the graphics content. At run time, sprites are re-rendered by priority within the budget of a single frame. Although the greedy sprite prioritization methods are promising, more comprehensive forms of optimization and control warrant study.

5 At the Interface of Computation and Cognition

Optimization of the flexible computational processes in Talisman poses an interesting set of problems at the interface between computation and cognition. As highlighted by Figure 6, we wish to move beyond information displayed in Figure 2 to map the links between perceptual quality and the metrics used to characterize the degradation of graphics along different dimensions (e.g., total squared distance deviation of mesh, tolerance of warp error-estimate, diminished resolution). We also need to characterize how various combinations of degradations, including those that can be controlled as part of the current Talisman architecture, can influence subjective impressions of the quality of a scene. After all, our goal is to provide content that is visually satisfying to people. Simple scientific goals focused on maximizing precision across the board are likely to be naive from the point of view of genuinely optimizing the final visual result.

The intertwining of computational strategies for rendering graphics and the nuances of the human perceptual apparatus allude to an analogous mix of the functioning of a graphics system and cognition with simpler technology. Thomas Edison, and others active in the early days of experimentation with the first motion pictures, discovered that the intrinsic latency and persistence of images in the human perceptual system could be exploited to generate the illusion of continuous experience when individual static frames were displayed in rapid succession. Today, we take for granted the well-known 0.1 second latency in the visual system and its implications for combining multiple still frames to generate the perception of objects in motion. Other, less known, effects have been harnessed in displays. For example, latencies in the visual system have been harnessed in sequential color systems for combining rapidly sequenced primary colors to generate arbitrary colors.

There is opportunity for exploiting perceptual insensitivity to minimize resource consumption in the new flexible Talisman “projector.” We need to understand sensitivities and blindspots to various kinds of degradations so as to ensure that our architectures and resource-allocation policies avoid or take maximal advantage of them. Studies of the human visual system within the domain of cognitive psychology can provide us with initial insights that can be later refined with more detailed studies of the specifics of Talisman approximations.

5.1 Relevant Findings in Cognitive Psychology

The cognitive psychology of vision is an active area of research fraught with competing and complementary theories, numerous studies, and an array of interesting results. Several areas of investigation in visual perception have potential relevance to our pursuit of links between perceptual quality of images and rendering approximations.

One paradigm for measuring the ability of human subjects to identify various features in scenes are based on a visual-search methodology. Studies of visual search have attempted to measure the abilities of human subjects to notice various components of scenes. The research has uncovered two distinct, but interrelated classes of visual processing, referred to as preattentive and attentive vision, respectively. *Preattentive* vision is thought to continually scan large areas at a time in parallel, efficiently noting features representing basic changes in pattern or motion. *Attentive* visual processes refer to the more serial, resource-limited processes found to be required to recognize details about objects and relationships in scenes.

Neisser noted that features efficiently detected by the preattentive visual processes include the overall color, size, luminance, motion, temporal onset of patterns, and simple aspects of shape like orientation, curvature (but not closure, gaps or terminators) [Neisser, 1963]. Julesz defined a class of features efficiently discriminated by preattentive vision, referred to as *textons* [Julesz, 1981]. Textons include elongated shapes such as ellipses, rectangles, and line segments that have specific colors, widths, lengths, orientations, depths, velocities, and flicker. Textons also include the ends of line segments, referred to as terminators, and the crossing of lines.

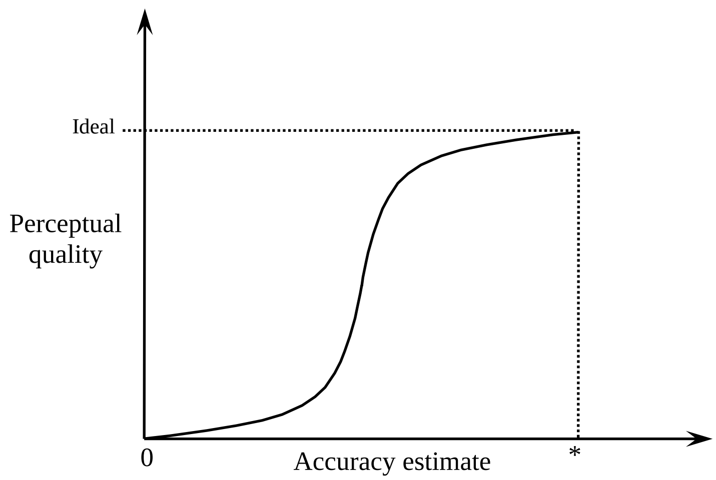

Figure 6: Relating measures of error to visual perception. The regulation of tradeoffs can be enhanced by employing knowledge about the relationships between common error measures and perceived changes in the quality of graphics.

Preattentive processing has been shown to be limited in its ability to detect the absence of features in a target amidst a sea of similar targets that contain the feature (*e.g.*, finding a circle without a slash among circles with a slash). More generally, the parallel, preattentive processes cannot efficiently identify cases where distinct features are conjoined into higher-level patterns or objects; identifying conjunctions of features requires the more focused, resource-strapped attentive vision processes.

Eriksen and Hoffman [Eriksen and Hoffman, 1972] demonstrated that attentive processing can be varied in scope and accuracy to explore small areas with high-resolution or be diffused over a wider area with loss of detail. Treisman and Gelade have studied the ability of people to recognize conjunctions of features. The team proposed and found evidence for the *feature-integration* theory of attention, where features are detected early on but are only related to one another to form recognizable objects with focused attention [Treisman and Gelade, 1980]. They also showed that recognition tasks were diminished by distraction and diverting of attention.

Several studies have further elucidated the links between preattentive and attentive processes. For example, researchers have found that objects may be recognized rapidly through efficient interactions of preattentive and attentive processes and search can be made more efficient through training. An example of efficient recognition of objects is the "pop out" effect, where objects seem to jump out of background patterns. Wolfe, *et al.* performed studies suggesting that serial search for conjunctions can be guided and made more efficient taking advantage of parallel processes [Wolfe et al., 1989]. The group proposed that preattentive processes can filter out distracters from candidates, and, thus, reduce the size of the serial search. This effect appears to depend on the quality of the guidance provided by the parallel processes and enhanced when elements are distinguished by luminance and color contrast, or when there are discontinuities in spatial or temporal patterning of the first-order properties giving rise to motion or texture differences [Cavanagh et al., 1990].

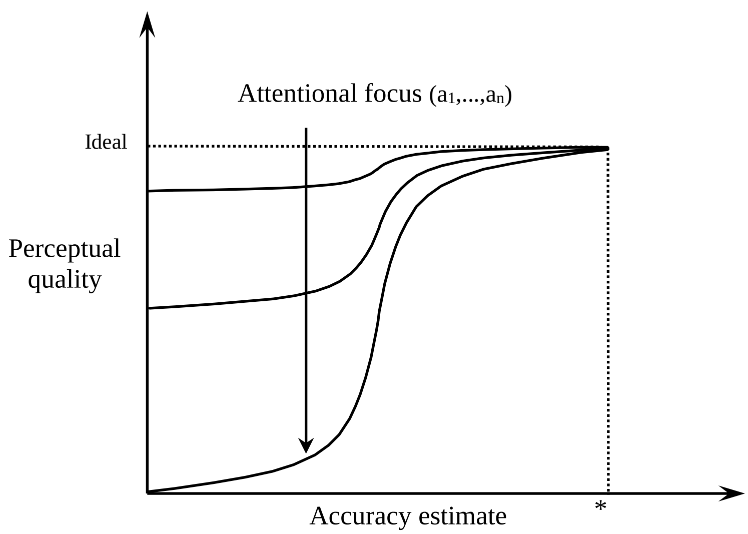

Figure 7: Attention and perception. Psychological studies of visual search in attentive vision suggest that a viewer's attention may have significant influence on the perception of degradation. We seek confirmation of such dependencies.

5.2 Implications of the Findings

To be sure, much of the work on visual search has centered on the study of reaction times and accuracy for recognition tasks with a goal of characterizing and discriminating different types of visual processing. It is not clear how the results apply to the likelihood or degree to which viewers will notice various spatial and temporal degradations produced by approximations in Talisman. Nevertheless, the results can guide our search for phenomena we may wish to study carefully in designing and controlling a flexible graphics rendering architecture.

In particular, results about the difficulty of viewers recognizing objects and missing details in objects in time-pressured search may have implications in a model of perceptual cost for image quality. For example, the finding that attentive vision is a serial, resource-constrained process suggests that the perceptual cost of introducing artifacts that can be recognized only via serial search is likely to be sensitive to the details of a viewer's attention. Such costs may be minimized when the viewer is focused on a different object or is distracted by the overall complexity of a scene.

Also, the findings characterize spatial and temporal features that can help guide attentive processes to recognize objects and patterns. These results suggest that artifacts generated via alternate degradations may be especially prominent if they lead to discontinuities in such attributes as color, texture, or motion. Thus, a set of multiple adjacent objects that have a similar structure but that have different textures—as may be the case when varying spatial resolutions are applied to similar adjacent objects—may lead to a “popping-out” of artifacts, highlighting the salience of the relative degradation. Likewise, it may be distracting, and, thus costly, to change the level of quality of an object or one or more sprites comprising an object over time—even when the changes center on transforming a low quality model or warp into a high-quality rendering as computational resources become available. A rich utility model could represent such information as the benefit of maintaining homogeneous quality for a subset of spatially interrelated sprites or objects.

Unfortunately, details of the perceptual costs of approximations provided by Talisman have not been addressed directly by the earlier psychological studies. Armed with knowledge gathered from the studies of human perception, we wish to ascertain these perceptual costs. Thus, we are collaborating with psychologists to perform studies of the perceived losses in quality associated with key Talisman approximations. We are interested in characterizing changes in perceptual quality along single as well as multiple dimensions of degradation. We are also pursuing answers to questions about the perceptual costs associated with heterogeneities in the quality at which interrelated sprites or objects are rendered over time and space.

5.3 Toward Models of Visual Attention

The results of experiments on visual search make us keenly interested in the influence of the viewer's focus of attention and the overall scene complexity on the perception of degradation in quality. As highlighted by the set of curves in Figure 7, we seek to understand the influence of attention on the functions linking perceptual quality to various approximations of rendering sprites or objects.

Given the results of the psychological studies, we put forth the conjecture that there will be a significant dependency between the perceptual costs of degradation and the focus of attention for artifacts that are processed by attentive vision, and that the overall complexity of a scene should slow a user's ability to identify artifacts in specific objects. We also believe that viewers may show significant toleration to the degradation of the resolution of objects with their apparent distance in foreground and with the motion of objects. Authors of graphics content have long employed these intuitions in the design of interactive systems.

The likely dependence of a viewer's focus of attention on the perceived costs of rendering approximations highlights opportunities for dynamically allocating resources based on attention. It may be valuable to construct coarse models of attention that identify sprites or objects that are likely to capture a viewer's focus of attention, based on a consideration of such attributes as story line, centrality of the object in the image, the z position, representing the object's depth in the scene, the motion of objects, the class of an object (*e.g.*, terrain versus character), and details about the spatiotemporal configuration of multiple objects. In particular, we seek models of attention that provide a probability distribution over sprites or objects that will be attended to given a pattern of evidence, E, including key attributes of a storyline and scene. The promise of characterizing the allocation of attention to scenes is captured by Figure 8.

Employing such models with deterministic measures of the perceived costs of degradations would allow us to compute an *expected cost* of a frame or set of frames. For example, if we have access to the probability that a viewer is attending to each of the contiguous objects, O_i, in a rendered scene, and assume that the cost of degradation of sprites comprising each object, S_{ij}, is zero when a viewer is not attending to that object, the overall expected cost associated with employing particular degradations D_k to render sprites, would be,

$$\text{Cost} = \sum_i \sum_j \; p(A^{S_{ij}} \mid A^{O_i}, E)\, p(A^{O_i} \mid E)\, C(D_k, S_{ij}) \qquad (2)$$

where $p(A^{O_i} \mid E)$ is the probability of a user attending to an object, $p(A^{S_{ij}} \mid A^{O_i}, E)$ is the probability that a user will attend to sprites *j* of object *i*, given that the viewer's attention is drawn to that object, and $C(D_k, S_{ij})$ is the perceptual cost of applying degradation strategy D_k to render sprite S_{ij}.

The task of building probabilistic models can be simplified by employing parameterized models of attention that generate likelihoods that artifacts in objects in a scene will be noticed by a viewer *given* a current *primary focus* of attention. Such parameterized models provide the probabilities that viewers will turn their attention to other objects as a function of the spatial and temporal relationships to the primary focus of attention, and the amount of time an artifact persists. The models of relative attention provide conditional probabilities of the form, $p(A^{O_i} \mid A^{O*}, E, t)$, where *t* is the amount of time an artifact persists, and A^{O*} is the object at the focus of attention. Such conditional probabilities are substituted for $p(A^{O_i} \mid E)$ in Equation 2.

Short of attempting to completely automate inference about focus of attention, such parameterized models of the likelihood of attention can allow authors of graphics content to provide inputs solely about the primary focus of attention in a sequence of frames, rather than deliberating about priorities of multiple sprites. When rendering the scenes, the system could assign costs to degrading sprites depending on their spatial and temporal relationships with objects at the focus of attention.

Summary

We have described key issues with our ongoing investigation of the control of computation for optimizing the quality of graphics created by a flexible graphics rendering system named Talisman. Talisman provides challenging problems with the control and optimization of flexible computational processes. The project requires us to focus on the details of the value structure of human visual perception and the relationships of perceived losses in the quality of images to computational savings achieved via the approximations. We seek to develop better control strategies by characterizing the interdependent costs of alternate phases of computation and by elucidating the particular sensitivities—and insensitivities—of human vision and attention. In particular, we hope more detailed models of computational and perceptual costs will allow us to achieve gains in tractability by selectively degrading rendered images so as to direct artifacts into the forgiving “blind spots” of the human visual system.

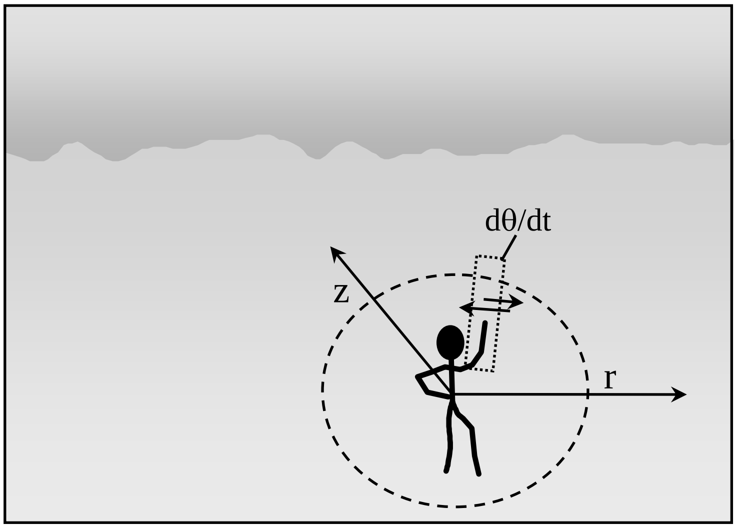

Figure 8. Role of visual attention. We seek to understand the variation of the sensitivity of viewers to various classes of graphics degradation as spatial and temporal relationships change. This figure highlights opportunities for modeling attention as functions of the primary focus of attention and such relationships as adjacency (*r*), distance in the background (*z*), and motion of sprites ($d\theta/dt$).

Acknowledgments

We thank Mary Czerwinksi and Kevin Larson for providing insights and pointers on psychological studies of visual search and for assistance with the design of experiments on visual perception and attention, Jim Kajiya for sharing his insights about tradeoffs and control of graphics rendering, Hugues Hoppe for input on model simplification strategies, Bobby Bodenheimer and Larry Ockene for their input on the details of the Talisman architecture, Ken Cooper, Robert Gross, and Channing Verbeck for input on tradeoffs and control of audio, and Mike Jones for ongoing discussion on links to work on resource-aware operating systems.

References

[Breese and Horvitz, 1990]
J.S. Breese and E.J. Horvitz.
Ideal reformulation of belief networks. In *Proceedings of Sixth Conference on Uncertainty in Artificial Intelligence, Cambridge, MA*}, pages 64–72. Association for Uncertainty in Artificial Intelligence, Mountain View, CA, July 1990.

[Cavanagh et al., 1990]
P. Cavanagh, M. Arguin, and A. Triesman. Effects of surface medium on visual search for orientation and size features. *Journal of Experimental Psychology: Human perception and Performance*, 16:479–491, 1990.

[Chen and Williams, 1993]
S.E. Chen and L.~Williams. View interpolation for image synthesis. In *Proceedings of SIGGRAPH 93*, pages 279–288. SIGGRAPH, August 1993.

[Dean and Wellman, 1991]
T.L. Dean and M.P. Wellman. *Planning and Control*, chapter 8.3 Temporally Flexible Inference, pages 353–363. Morgan Kaufmann Publishers, San Mateo, California, 1991.

[Eriksen and Hoffman, 1972]
C.W.W. Eriksen and J.E. Hoffman. Temporal and spatial characteristics of selective encoding from visual displays. *Perception and Psychophysics*, 12:201–204, 1972.

[Funkhouser and Sequin, 1993].
T.A. Funkhouser and C.H. Sequin. Adaptive display algorithm for interactive frame rates during visualization of complex virtual environments. In *Proceedings of SIGGRAPH 93*, pages 247–254. SIGGRAPH, August 1993.

[Hoppe, 1996]
H. Hoppe. Progressive meshes. In *Proceedings of SIGGRAPH 96*, pages 99−108. SIGGRAPH, August 1996.

[Horvitz, 1987]
E.J. Horvitz. Reasoning about beliefs and actions under computational resource constraints. In *Proceedings of Third Workshop on Uncertainty in Artificial Intelligence*, pages 429–444, Seattle, Washington, July 1987. AAAI and Association for Uncertainty in Artificial Intelligence, Mountain View, CA. Also in L. Kanal, T. Levitt, and J. Lemmer, ed., *Uncertainty in Artificial Intelligence 3*, Elsevier, Amsterdam, 1989, pp. 301-324.

[Horvitz, 1988]
E.J. Horvitz. Reasoning under varying and uncertain resource constraints. In *Proceedings AAAI-88 Seventh National Conference on Artificial Intelligence, Minneapolis, MN*, pages 111–116. Morgan Kaufmann, San Mateo, CA, August 1988.

[Horvitz, 1990]
E.J. Horvitz. *Computation and Action Under Bounded Resources*. PhD thesis, Stanford University, 1990.

[Horvitz and Breese, 1990]
E.J. Horvitz and J.S. Breese. Ideal partition of resources for metareasoning. Technical report, Knowledge Systems Laboratory, Stanford University, February 1990. KSL-90-26.

[Julesz, 1981]
B. Julesz. A theory of preattentive texture discrimination based on first order statistics of textons. *Biological Cybernetics*, 41:131–138, 1975.

[McMillan and Bishop, 1995]
L. McMillan and G. Bishop. Plenoptic modeling: an image-based rendering system. In *Proceedings of SIGGRAPH 95*, pages 39–46. SIGGRAPH, August 1995.

[Neisser, 1963]
U. Neisser. Decision time without reaction time: Experiments in visual scanning. *American Journal of Psychology*, 76:376–383, 1980.

[Regan and Pose, 1994]
M. Regan and R. Pose.
Priority rendering with a virtual reality address recalculation pipeline. In *Proceedings of SIGGRAPH 94*, pages 155–162. SIGGRAPH, August 1994.

[Shade96 et al., 1996]
J. Shade, D. Lischinski, D.H. Salesin, T. DeRose, and J. Snyder. Hierarchical image caching for accelerated walkthroughs of complex environments. In *Proceedings of SIGGRAPH 96*}, pages 75–82. SIGGRAPH, August 1996.
[Torborg and Kajiya, 1996]
J. Torborg and J. Kajiya. Talisman: Commodity realtime 3D graphics for the PC. In *Proceedings of SIGGRAPH 96*, pages 353–363. SIGGRAPH, August 1996.

[Treisman and Gelade, 1980]
A.M. Triesman and G. Gelade. A feature-integration theory of attention. *Cognitive Psychology*, 12:97–136, 1980.

[Wolfe et al., 1989]
J.M. Wolfe, K.R. Cave, and S.L. Franzel.
Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology*, 15(3):419–433, 1989.

[Zilberstein, 1993]
S. Zilberstein. *Operational Rationality through the Compilation of Anytime Algorithms.* PhD thesis, Computer Science Division, University of California, Berkeley, 1993.