Technical
Brief

The Key to a Stunning Interactive 3D Experience:
High Performance + Image Quality

NVIDIA

The Key to a Stunning Interactive 3D Experience:

High Performance + Image Quality

Executive Summary

A realistic 3D experience is possible only through a combination of high performance and high quality image rendering. *While a frame rate of 30 to 60 frames per second is necessary to see smooth 3D motion, the quality of the models in each of those frames is the real key to a stunning 3D experience.*

Creating perfectly realistic image quality is an infinitely difficult problem. *There are a myriad of features that contribute to image quality, including higher polygon counts, sophisticated per-pixel shaders, cube environment mapping and more.*

NVIDIA's image quality is second to none. *NVIDIA's GeForce2 family of GPUs offers the widest array of features to enhance the image quality of current and future applications, all with superb performance and compatibility.*

Hardware transform and lighting provides rich environments and highly realistic characters. *The CPU does game play, physics and artificial intelligence calculations. The GPU does the transform, lighting, and rendering calculations. This perfectly balanced model of graphics processing provides an optimal computing environment for the richest, most realistic 3D experience.*

Per-pixel shaders make sophisticated surface and material rendering possible. *Sophisticated lighting effects, realistic bump mapping, and other seriously cool effects become possible with NVIDIA's per-pixel shading technology. Wood looks like wood, satin looks like satin and metal looks like metal.*

Cube environment mapping allows realistic reflections and other effects*.* *Unlike traditional reflection mapping techniques, cube environment mapping dynamically maps the surrounding environment to the object perfectly.*

Introduction

Today's baseline 3D graphics processors (such as the NVIDIA TNT2™) have, at a minimum, image quality features such as Gouraud shading, 32-bit color, bilinear filtering, and trilinear mip-mapping. That means today, an average, low-cost consumer desktop has 3D hardware capability rivaling the workstations of a few years ago.

However, these baseline features fall far short of the goal of photorealistic rendering at 60 frames per second. To achieve that goal, the graphics pipeline has to include even more features while *simultaneously* increasing performance. While we are several generations away from achieving truly interactive, photorealistic image quality, modern GPUs are far more advanced than those of even one year ago. This paper examines the complex topic of image quality in 3D rendering.

There is no single, magical feature that yields optimum image quality. Instead, superb image quality requires a number of features, in addition to good performance, to achieve a visually stunning result. All of these features need to be able to work as a whole, so that turning on one feature doesn't adversely affect the performance or capability of another. It's not an easy task.

The GeForce2 family of GPUs offers the most advanced image quality architecture in the world today. Every member of the family supports a wide range of features, including:

- Hardware transform and lighting (T&L)
- Per-pixel shading effects
- Cube environment mapping
- Superior performance at high resolutions
- Superior performance over competing solutions

Let's take a closer look at the features of the GeForce2 family of GPUs and see how each contributes to stunning, interactive 3D.

Transform and Lighting

Transforms are mathematical calculations used to determine changes in the location, size, shape, or orientation of a geometric object as it's being displayed on the two dimensional display of a PC. They are computationally very expensive, and require a huge number of floating point calculations.

NVIDIA's hardware transform and lighting capability has two purposes:

- Enable the creation of richer experiences by allowing developers and artists to create complex environments, more realistic characters, and smoother animation.
- Offload transform and lighting calculations from the CPU, to enable smarter AI, more robust collision detection, better pathing algorithms, and highly realistic physics.

The GeForce2 GTS is capable of peak geometry throughput of 25 million triangles per second; the GeForce2 MX can pump out 20 million triangles per second. The overall geometry performance is roughly 3x the best performance available in high-end workstations of only two years ago. To facilitate offloading the CPU, all vertex data is stored in AGP memory, so when the geometry data is pulled into the GPU, the CPU doesn't have to manage the transfers.

If software transform and lighting is used, the CPU has to handle T&L and all other gameplay elements, so there is a constant juggling act between various game elements, including graphics -- even with a very fast CPU. Since even fast CPUs can't match the geometry performance of NVIDIA's GPU technology, severe compromises would have to be made. The frame rate may vary radically as the CPU switches between geometry, AI, and other key game functions. Offloading the transform and lighting to the GPU frees up the CPU and creates a much smoother experience for the player.

Currently, developers view T&L as a way to boost performance. However, as they become more comfortable with GPU technology, and as products like NVIDIA's GeForce2 MX continue to push the performance curve, game designers will expand their use of T&L to create highly realistic environments and characters by increasing polygon counts. The result will be games that look better than ever before.

With T&L, curves are rounder, character joints look more natural, and the overall look of the game is much richer and more realistic. In this case, the visual quality of the title greatly benefits, all without dropping below acceptable frame rates.

Here's a screenshot from Shiny's upcoming title, *Sacrifice*. Notice the huge polygon counts on the screen.

Figure 1. Note the smooth, rounded nature of the characters in this shot from Sacrifice*. The smoothness is made possible by the high polygon count, as seen in the right hand image.*

NVIDIA's GeForce2 family makes these smooth images possible by allowing the developers to use a higher polygon count in their characters. The T&L engine on the GeForce2 family is the key enabler for these high-polygon-count characters.

Hardware transform and lighting, coupled with other features of the GeForce2 family, offer a visual experience unlike anything that existed before on the PC desktop. For example, look at the cave and water in the following image rendered in real time on a GeForce2 MX.

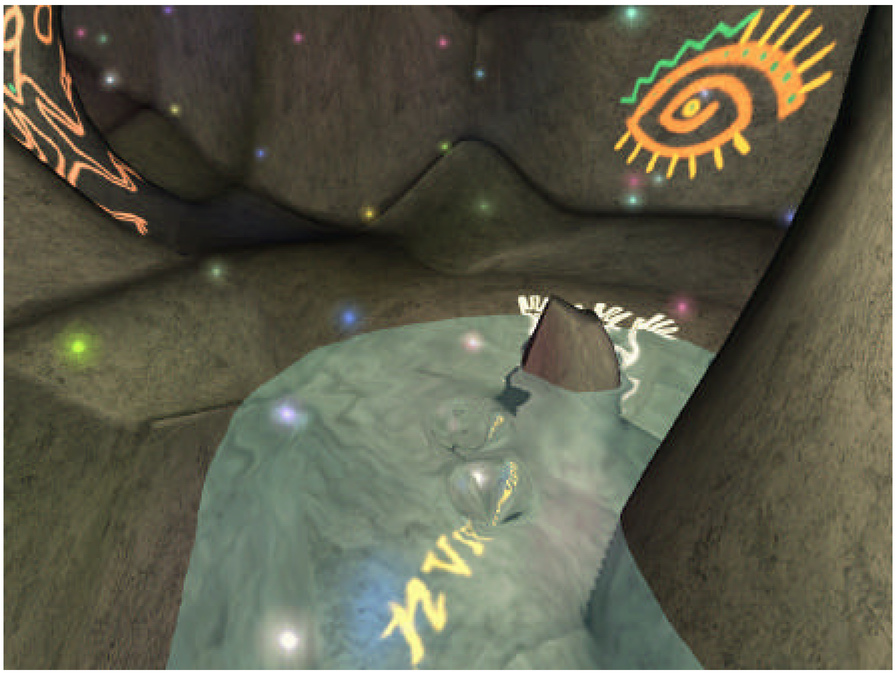

Figure 2. Here's a cave pool image, rendered with reflection, refraction, and transparency.

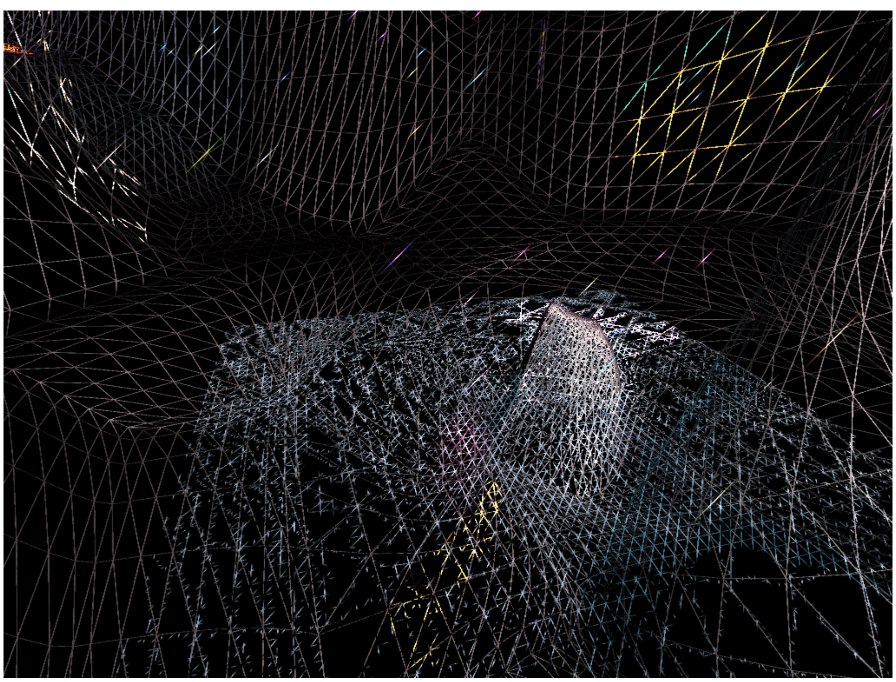

Figure 3. The same scene, in wireframe mode, shows the large polygon count needed to make good-looking 3D water.

NVIDIA Shading Rasterizer (NSR)

The NVIDIA Shading Rasterizer is another key step along the path to the ultimate goal of high-performance, photorealistic animation, and is an important part of NVIDIA's image quality architecture. The NVIDIA Shading Rasterizer, as implemented in the GeForce2 family, gives application developers the freedom to create sophisticated, per-pixel effects, ranging from highly accurate lighting and shadows to bump mapping that is so real, you can almost feel the texture under your fingertips.

Per-Pixel Lighting

The GeForce2 family is the first line of mainstream consumer grade GPUs to offer real-time, per-pixel lighting. They provide hardware support for per-pixel operations by using register combiner functionality. This enables a texture blending operation called a dot product that enables per-pixel calculations. Available to developers in Microsoft® DirectX® 6 and above (D3DTOP_DOTPRODUCT3) as well as OpenGL (via the NV_register_combiners extension), dot product texture blending operations allow diffuse, specular, spot, and point light effects to be calculated dynamically on a per-pixel basis.

Take a look at the following images from Computer Artworks' recently shipped title, *Evolva* rendered in real time on a GeForce2 MX.

Figure 4: The NVIDIA Shading Rasterizer in action. Note the realistic, rough surfaces on the creature and terrain. This is made possible by the NVIDIA Shading Rasterizer.

Without the per-pixel capability, the same scene looks flat and much more artificial:

Figure 5: The same scene, rendered without per-pixel effects. The creature doesn't stand out from the terrain as well, and the flat ground looks artificial.

Real-time, per-pixel shading effects offer a host of capabilities previously unavailable on any graphics hardware product in the consumer market. Per-pixel effects can be easily added to existing game engines – allowing developers to utilize a number of interesting effects. The first is per-pixel bump mapping – a technique in which a texture (the bump map) is applied to an object to make the surface appear to be rough. GeForce2 family's per-pixel bump mapping remains accurate and realistic even as light moves across the surface.

Figure 6: As the light source moves over the object, the shadows cast by the bumpy surface remain accurate. Also, note that this doesn't look artificially shiny, unlike environmental bump mapping.

Per-pixel lighting is also useful when you have large, flat surfaces. Big, flat objects, like walls and floors, often consist of very few triangles. Per-pixel lighting can look terrific on large, flat objects that consist of such limited geometry. Couple per-pixel lighting with per-pixel bump mapping and you can create powerful imagery with relatively low triangle counts.

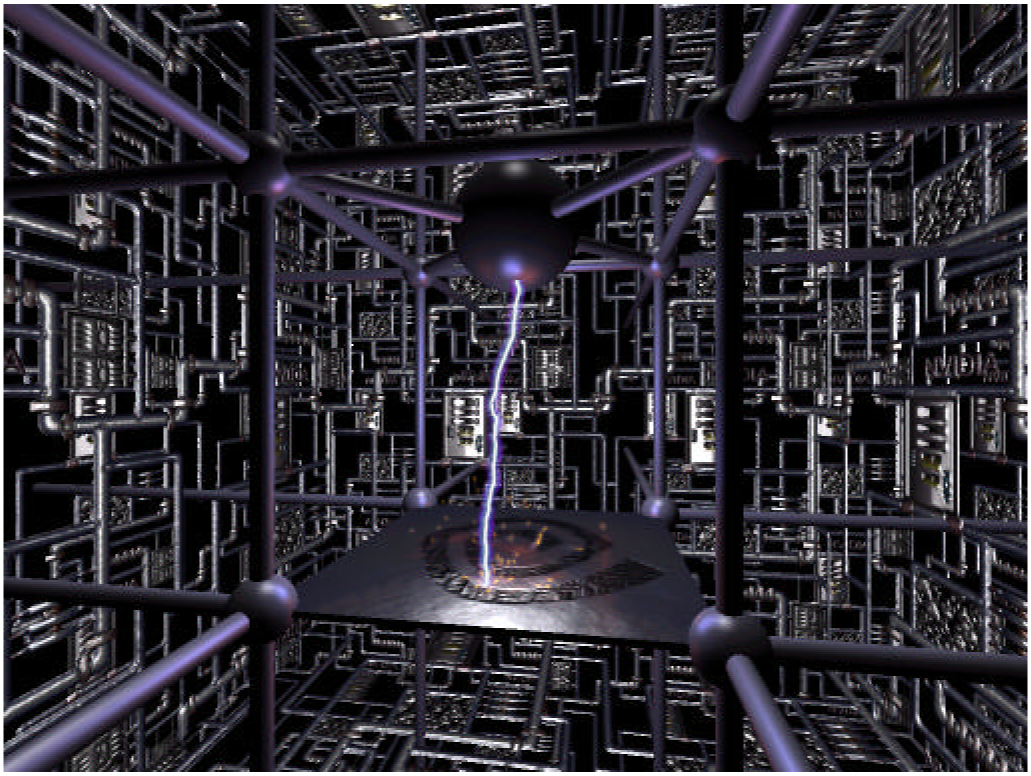

Figure 7: In this screen shot, the particles are actually per-vertex hardware lights, whereas the background lighting is per-pixel shading effects using texture maps.

Cube Environment Mapping

In the past, reflections or shiny surfaces in 3D applications have looked inaccurate and artificial. Creating accurate reflections is a tough challenge, and the older, spherical mapping technique often looked unrealistic. Cube environment mapping creates accurate reflections by using six different maps, rather than the single map used by spherical reflection maps.

Reflection mapping is often used in arcade driving games, but is widely used in other titles. Here's an image from Codemaster's upcoming PC release of *Colin McCrae Rally 2*.

Figure 8: Cube environment mapping is responsible for the realistic reflections on the rally car. Incidentally, this car consists of over 700 polygons.

High Resolutions

The best way to achieve clean lines and crisp visuals is to run applications at very high resolutions. More pixels rendered means smoother lines. The ability to run at 1280 x 1024 or even 1600 x 1200 in many titles has brought stunning image quality to the consumer desktop. High-resolution gaming complements high polygon counts superbly, and the combination will make for an unparalleled gaming experience.

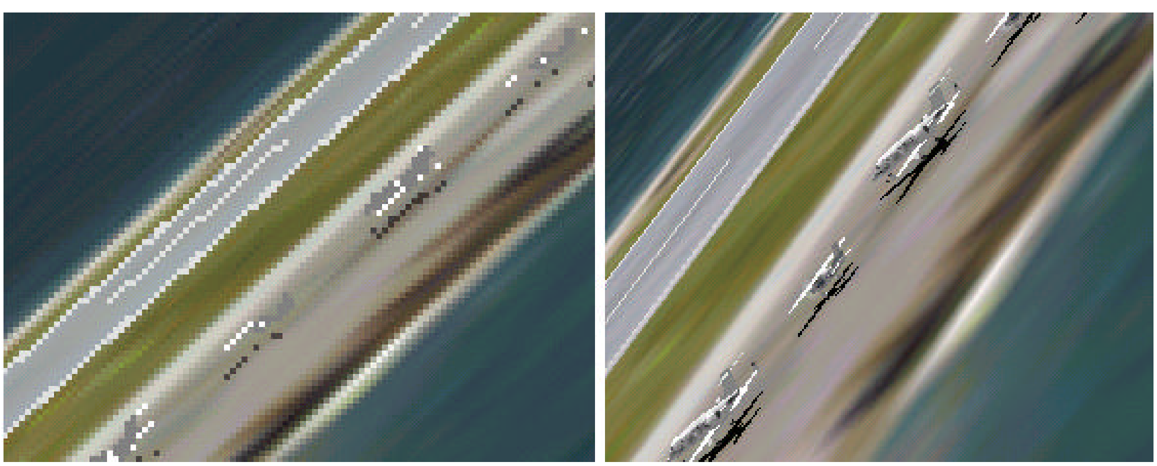

Figure 9: The screen shot on the left was taken at 640 x 480; the right-hand shot is 1600 x 1200. Note that the planes on the right are recognizable as aircraft on the ground, not just a blob of pixels.

High resolution adds crisp definitions to details. Flight simulation enthusiasts can see enemy aircraft farther away than at lower resolutions. *Quake III* players can see enemies farther down a long corridor than before. Strategy gamers can now see more of the map than ever before. Graphics professionals can now run CAD and modeling applications at much higher resolutions, including real-time rendering previews.

With fast graphics chips, high performance CPUs, inexpensive 19" monitors, and most standard frame buffers exceeding 16MB, there's no reason not to run game titles at higher resolutions. It's the natural evolution of computer graphics capabilities.

Color and Depth Precision – The Higher the Better!

Our real world includes objects comprised of an infinite number of shades of all colors of the spectrum. In addition, the position of objects relative to one another can be measured in infinitely small units. Therefore, accurate rendering requires a great deal of precision in both color and Z-buffer. (The Z-buffer is used to position the depth of 3D objects relative to one another). 32 bits of color and 32 bits of Z/stencil are absolutely necessary in modern GPUs. These are fundamental features of all NVIDIA products.

Performance – The Final Piece of the Formula

Having terrific image quality features is pointless if the graphics technology can't deliver the performance needed for today's high performance, real-time 3D applications. NVIDIA's GeForce2 family is capable of delivering frame rates in excess of 60 frames per second at high resolutions in most of today's 3D applications. As more games support hardware transform and lighting, gamers will experience richly detailed characters and environments with the blistering performance that's become the hallmark of NVIDIA GPUs.

Conclusion

In the end, only the GeForce2 family delivers the combination of performance and image quality needed for both today's game titles and tomorrow's richly detailed, virtual worlds. Through a combination of hardware transform and lighting, per-pixel shading, and cube environment mapping, users will be able to engage in immersive experiences never before seen on a desktop PC. The GeForce2 family has the performance for today's best applications and the features to keep end users happy for years to come.