Technical Brief

CineFX 4.0

Unimaginable Speeds for Advanced Visual Effects

Building In More Power and Visual Quality

The fourth-generation NVIDIA® CineFX® engine builds unimaginable speed into the NVIDIA GeForce® graphics processing units (GPUs). Using the CineFX 4.0 engine, developers can create and display the most advanced and high-quality visual effects for the hottest PC games (Figure 1) and other cutting-edge visual applications.

Every requirement for 3D visualization falls into one of two categories—performance or image quality—with the ultimate goal of being able to carry out more calculations in less time with the highest possible image quality.

The new NVIDIA GeForce 7 Series GPUs featuring the CineFX 4.0 engine incorporate architectural advancements that accelerate the most common operations required for 3D visualizations. This allows for more complex shader effects while maintaining the highest levels of image quality. The new design introduces innovation at every stage of the pipeline:

- A redesigned vertex shader unit reduces the time to set up and perform geometry processing.
- A new pixel shader unit design can carry out twice as many floating-point operations and greatly accelerate other mathematical operations to increase throughput.
- An advanced texture unit incorporates new hardware algorithms and better caching to speed filtering and blending operations.

Figure 1. NVIDIA Luna Technology Demo Rendered by GeForce 7800

By pushing every component of 3D visualization to new levels of performance and quality, the CineFX 4.0 engine lets users experience more realism in every frame. Existing applications will perform at jaw-dropping speeds with enhanced visual clarity and sharpness. And programmers will take advantage of the advanced processing capabilities to introduce new effects, increase the complexity of visualizations, and raise the level of the overall user experience—all without compromising on performance.

Vertex Shader

A Backbone of 3D Graphics

The vertex shader is one of the backbones of 3D graphics. 3D worlds are made up of models in which the basic graphical element is the triangle (Figure 2). Each corner of the triangle (known as a vertex) is referenced by a location in the 3D world, and is further described using various properties.

Figure 2. Wire Frame Model of NVIDIA Luna Technology Demo

The first stage of the vertex shader pipeline processes the individual triangle vertices. Each vertex is processed by looking at its location in the 3D world, as well as at the location of the camera and light sources in the world. Using 3D matrix transform calculations, this vertex information is then transformed into the 3D space of the camera.

Faster Triangle Setup

To prepare the newly transformed vertices for the next stage of rendering—which is carried out on a pixel-by-pixel basis—the vertices are grouped into triangles by the triangle setup unit (Figure 3).

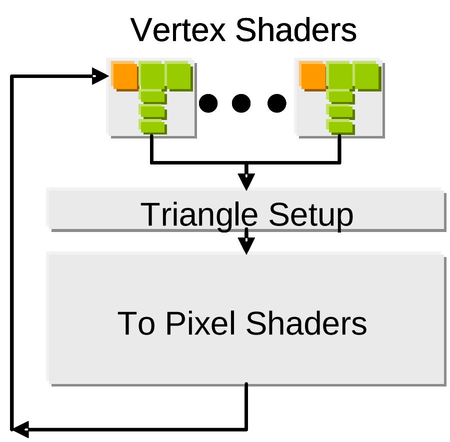

Figure 3. Vertex Shaders Feeding Data into the Triangle Setup Unit

Note: The GeForce 7800 GPUs have up to eight vertex shaders.

The triangle setup unit mathematically determines the "walk across" or *raster* of the triangle based on the three vertices of the triangle. During this process, the triangle is split into a number of raster lines. Color fragments are shaded as each line is rasterized. This algorithm continues extracting information until the entire triangle is filled with shaded fragment values (Figure 4).

Note: A fragment can be a pixel, and in most cases the two terms can be used interchangeably.

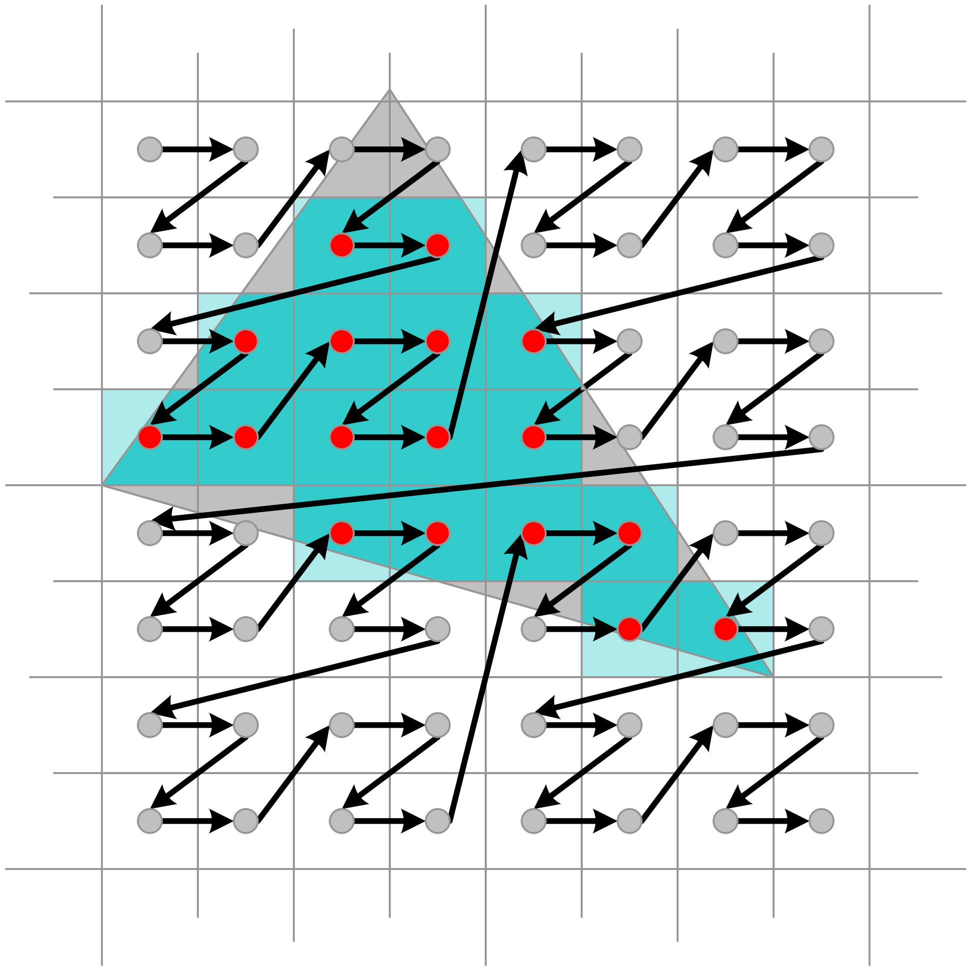

Figure 4. A Triangle Being Rasterized

Note: The rasterization pattern in Figure 4 is conceptual and does not represent the actual raster pattern of the GeForce 7800 GPUs.

Accelerating the triangle setup increases the overall throughput of the 3D pipeline. This is especially true in geometry- or vertex-bound applications such as shadow rendering.

Pixel Shader

Pixel shading is inherently math-intensive; graphics equations, for lighting in particular, are very complicated and computationally intense. Effects like refraction, reflection, normalization, and embossing require a large number of various mathematical calculations.

In addition, simulating complex materials in real-time requires a huge amount of computational horsepower.

Accelerating MADD Computations

Multiply and *accumulate* are commonly used math functions in 3D graphics. Also referred to as *multiply-add* (MADD) operations, they show up in transformations, lighting, normal map calculations, and many other operations.

The CineFX 4.0 engine accelerates MADD operations for overall increased throughput for the pixel shader. In fact, the GeForce 7800 can perform up to twice the MADDs of the previous-generation GPUs.

The relief mapping example illustrates the significance of the MADD function, and the following sections detail the performance improvements that result from this and the other innovations of the latest generation of NVIDIA GPUs.

Example: Relief Mapping

In relief mapping applications, a normal map that contains elevation data can be used to create the illusion of depth and height. In Figure 5, notice the ridges on the teapot and the objects emerging from the wood platter, both created using a special technique called *relief mapping*.

Figure 5. Relief Maps Used to Create Ridges on the Teapot and Raised Shapes on the Wooden Tray

The actual model in Figure 5 contains no geometry to create the bumps on the wood or the teapot. The relief information is stored as completely separate textures that can be applied to triangles for the desired look (Figure 6 and Figure 7). The depth or height is derived as a function of the base texture, the relief map, and scene lighting.

Figure 6. Brick and Wood Textures

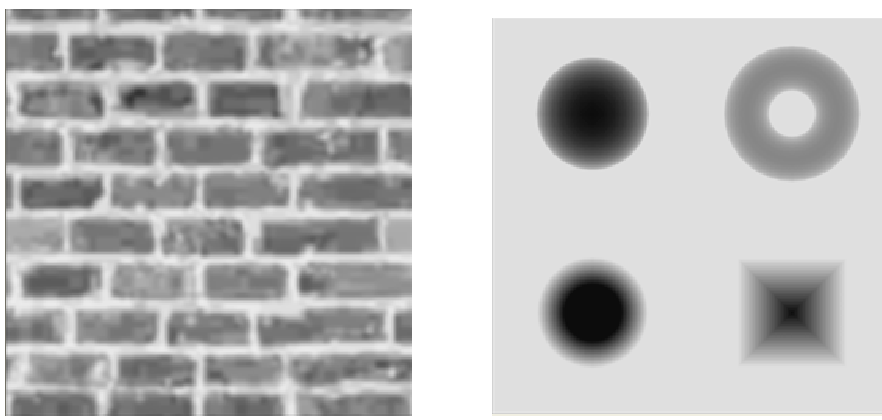

Figure 7. Relief Maps Create the Appearance of Depth When Combined with a Base Texture

To understand the impact of accelerated math functions, you need to be familiar with the required math operations. For calculating the types of 3D effects shown in Figure 5 the main equations for the pixel shader include code like this (Figure 8):

```
float4 main_rm()
{
      A = pixel position in global space (passed in to shader
in texcoord0)
      Viewdir = normalize( A – camera_pos );
      size = relief_depth / dot( -relief_normal, Viewdir);
      B = A + Viewdir * size;
      V = B – A;
      d1 = ray_intersect_rm(A, V);
      P = A + d1 * V;
      Lightdir = normalize( P – light_pos );
      size = relief_depth / dot( -relief_normal, Lightdir);
      C = P – d1 * size * Lightdir;
      D = C + Lightdir * size;
      V = D – C;
      d2 = ray_intersect_rm( C, V );
      if (d2<d1)
            // pixel in shadow
            color = shadow_color();
      else
            // apply lighting
            color = phong_lighting();
      return color;
}
```

Figure 8. Example of High-Level Shader Code

The assembly code that these types of shaders create is shown below in Figure 9.

Note: The code example in Figure 8 is only a small piece of the relief map shader program. For the full listing, please visit the NVIDIA Developer Web site at http://developer.nvidia.com.

```
texld r0, r0, s1
texld r1, r1, s1
add r0.z, -r1.w, c8.z
cmp r0.y, r0.z, c8.z, r3.w
add r0.z, -r3.w, c6.x
cmp r1.w, r0.z, r3.w, r0.y
add r0.w, -r0.w, c8.w
cmp r0.z, r0.w, c8.w, r1.w
add r0.w, -r1.w, c6.x
cmp r3.w, r0.w, r1.w, r0.z
mul r0, r2.xyxy, c9.xxyy
mad r1.xy, t0, c0.x, r0
mad r0.xy, t0, c0.x, r0.zwzw
mad r1.xyz, r4.x, t2, -r1
dp2add r0.w, r4, -r4, c5.w
add r2.xyz, r2, -t5
rsq r0.w, r0.w
dp3 r1.w, r2, r2
rcp r0.w, r0.w
rsq r1.w, r1.w
```

Figure 9. Relief Map Shader Code Snippet

The GeForce 7800 GPUs introduce accelerated math functionality that dramatically improves performance for the types of programs like the example in Figure 9, which include numerous MADD operations. The advanced NVIDIA architecture features speed and flexibility, letting developers introduce more complex effects that can be rendered at lightning-fast speeds.

Texture Engine

CineFX 4.0 includes a redesigned texture processing engine. Textures can be grabbed and accessed faster, and developers can take advantage of a variety of texel sample sizes. These improvements deliver major benefits for high-precision texture applications such as high dynamic-range (HDR) rendering. Anisotropic filtering also benefits from the improved cache design of the latest texture engine. Notice the high dynamic-range effects of the Unreal Engine 3 image in Figure 10.

Image Courtesy of Epic Games.

Figure 10. High Dynamic-Range Effects—Vibrant Brights and Darks—Are Only Possible with the 64-Bit HDR Capabilities of GeForce 7800

Combined with accelerated math functionality, beautiful HDR effects can be rendered with an extremely high level of performance and precision.

Other Advances

New Antialiasing Modes

Engineers at NVIDIA have created two new antialiasing modes—transparency adaptive supersampling and transparency adaptive multisampling. Both modes increase the quality and performance of antialiasing.

Transparency Adaptive Supersampling

Transparency adaptive supersampling and multisampling take additional texel samples and antialiasing passes to enhance the quality of thin-lined objects such as chain link fences, trees, and vegetation.

These types of objects are generally rendered on very simple polygon models (or even one polygon). The complexity of the final image (a group of branches or vegetation) comes from the texture that is mapped onto the polygon. Conventional antialiasing does not help this situation, because the edges of the vegetation or branches are actually inside the projected texture. Pixels inside a polygon are not touched by current antialiasing methods.

Transparency adaptive supersampling methods solve this problem by keying off information embedded in the alpha channel of the texture. Areas that have the key set can receive antialiasing even though they are not on the edge of a triangle. The result is a smoother, more beautiful image.

Figure 11 through Figure 13 demonstrate the advantage of this method.

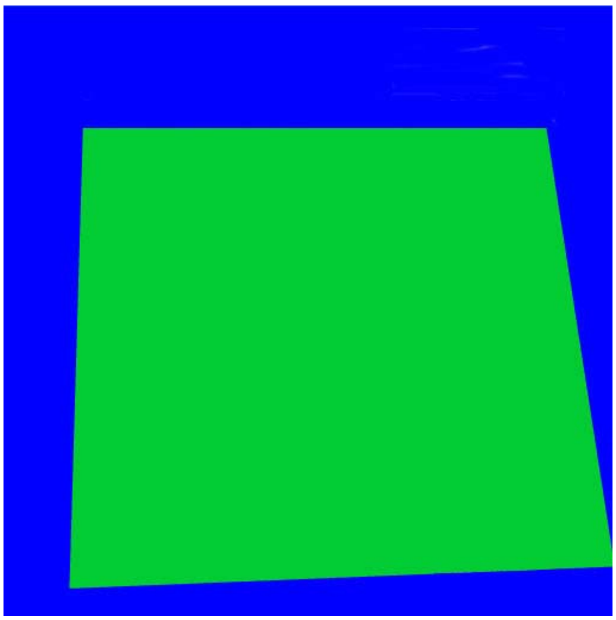

Figure 11. Single Polygon with a Single Blank Texture Rendered onto It

In Figure 12, notice how the edges are aliased and jaggy. Normal antiliasing would not improve the quality of the displayed image using this texture.

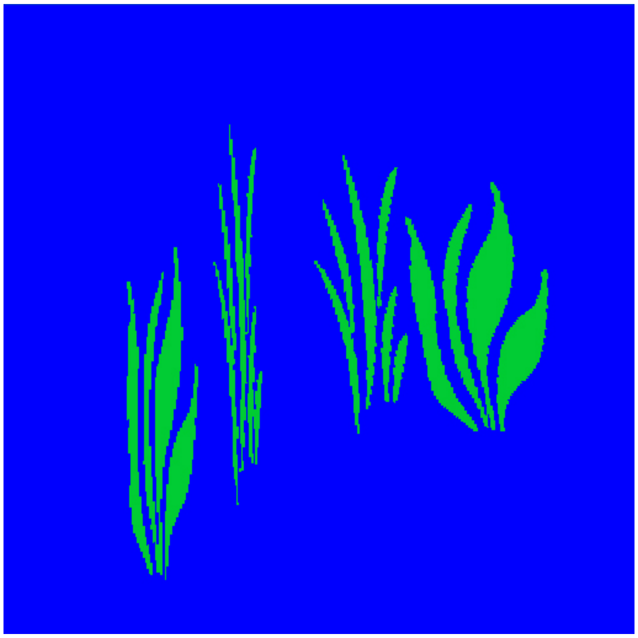

Figure 12. Same Single Polygon with a Single Texture Containing Plants Rendered into It

The example in Figure 13 has transparent sampling applied to the image. Notice how the blades of vegetation are now smooth and appealing.

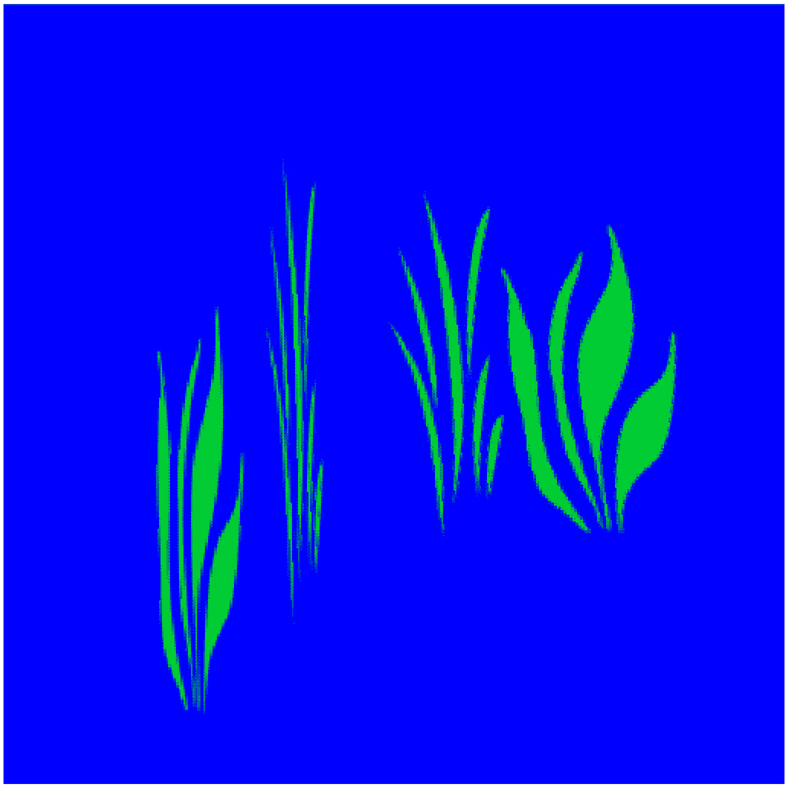

Figure 13. Same Texture Rendered into the Single Polygon

In general, supersampling is prohibitively expensive in terms of execution time. However given the NVIDIA CineFX 4.0 engine's innovative use of these adaptive algorithms, supersampling is only applied to selective parts of the image. The result is improved quality with good levels of performance.

Transparency Adaptive Multisampling

Transparency adaptive multisampling also improves antialiasing quality—with even higher levels of performance because one texel sample is used to calculate surrounding subpixel values. Although transparency adaptive multisampling is not as high quality as the supersampling method, its increased efficiency balances improved image quality and high levels of performance.

The benefit of this approach is shown in Figure 14. The visual improvements of adaptive supersampling are obvious when compared to generic supersampling/multisampling approaches. In particular, notice the detail in the tree branches and blades of grass.

Figure 14. Normal Antialiasing (Left) vs. Transparent Adaptive Antialiasing with Clear, Realistic, Small Branches and Foliage (Right)

Both new antialiaing modes—transparency adaptive supersampling and transparency adaptive multisampling—are ideal for outdoor environments, scenes with vegetation, chain link fences, and any situation where the models are seen as very thin from the angle of viewing.

Support for the Next Generation of Microsoft Operating System

The NVIDIA CineFX 4.0 engine supports the soon-to-be-available Microsoft® Longhorn operation system and the Microsoft Windows® Graphics Foundation 1.0 standard. Powered by the composited desktop hardware engine, many graphics advancements will be available right on the desktop.

These advancements include:

- Video post-processing
- Real-time desktop compositing
- Seamless multiple 3D applications
- Accelerated antialiased text rendering
- Special effects and animation

All of these features add value to the overall end-user experience. And through the desktop hardware composite engine, there is minimal context switching and no effect on CPU performance.

Conclusion

The NVIDIA CineFX 4.0 engine injects breakthrough graphics technology into the core levels of the vertex shader, pixel shader, and texture engines.

By accelerating triangle setup, crucial math elements of the pixel shader, and texture manipulations, the newest engine lets 3D graphics developers achieve new levels of performance and visual quality.

References

Oliveira, Manuel M., Gary Bishop, David McAllister. *Relief Texture Mapping. Proceedings of* ***SIGGRAPH 2000*** (New Orleans, LA), July 23-28, 2000.

Policarpo, Fábio, Oliveira, Manuel M., Comba, João L. D. "Real-Time Relief Mapping on Arbitrary Polygonal Surfaces." Proceedings of the 2005 Symposium on Interactive 3D Graphics and Games, pages 155 - 162.

http://www.inf.ufrgs.br/~oliveira/RTM.html

Notice

ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, “MATERIALS”) ARE BEING PROVIDED “AS IS.” NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE.

Information furnished is believed to be accurate and reliable. However, NVIDIA Corporation assumes no responsibility for the consequences of use of such information or for any infringement of patents or other rights of third parties that may result from its use. No license is granted by implication or otherwise under any patent or patent rights of NVIDIA Corporation. Specifications mentioned in this publication are subject to change without notice. This publication supersedes and replaces all information previously supplied. NVIDIA Corporation products are not authorized for use as critical components in life support devices or systems without express written approval of NVIDIA Corporation.

Trademarks

NVIDIA, the NVIDIA logo, CineFX, and GeForce are trademarks of NVIDIA Corporation.

Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright

NVIDIA Corporation
2701 San Tomas Expressway
Santa Clara, CA 95050
www.nvidia.com