NVIDIA™

Hardware Transform and Lighting

Richard Huddy

NVIDIA Corporation

RichardH@nvidia.com

A few general remarks…

- **Differences between H/W ‘T’ and S/W ‘T’:**
  - **Use the other HAL**
  - **Put your VBs where the driver wants them and not into sysmem**
  - **Can’t use optimized sysmem VBs to render**
  - **The output of H/W is pixels, *not* transformed vertex data**
  - **Optimize should (but might not be) be a no-op**

NVIDIA

What you can use it for…

- Rendering
- Generating textures (SRT)
- Multi-matrix blending
- High performance vertex lighting

NVIDIA.

What you can't use it for…

- Getting transformed data back
- Occlusion testing
- Sort independent alpha ☹
- Some multi-pass fogged effects

NVIDIA

Mixing and matching

- Multi-bone blending
  - You do the blending in object space
  - Leave final transform and lighting to us
- “Load balancing”
  - Usually not a great idea but can work
- Multi-pass rendering with problems (like pure modulate).

NVIDIA.

Avoiding mixing

- Good:
  - Using the fog table to avoid S/W transform with multi-pass effects
  - Doing cunning things with fog values
- Bad:
  - Using ProcessVertices to ‘help out’ - it’s likely to slow you down.

ProcessVertices is…

- Capable of generating post-transform data
- Always executed on the CPU
- Great for bounding boxes
- Fine for generating ‘static’ TLVertex data
- Not worth using on TnL devices unless you have a massive computational per-vertex load

NVIDIA.

H/W TnL is for what data?

- Static
  - Put it into VBs
  - Use the create/lock flags informatively
  - Optimize it if it's going to be here long
- Dynamic
  - Put it in VBs
  - Manage your VBs carefully
  - Don't optimize it

Culling and Clipping

- Do gross culling on the CPU
- Expect H/W clipping to be fast (GeForce adds no extra cost for clipping)
- Expect guard band clipping to be v. fast
- Don’t cull individual polys unless you cull them very early and they are v. expensive
- H/W will clip out $1.0 < z < 0.0$

How fast is the transform?

- Typically expect it to be limited by the memory speed
- This implies that untransformed is not slower than transformed data
- 25 Million vertices per second?

NVIDIA

TnL HAL Lighting capabilities

- **Maximum number of lights active at any one time (very different from S/W)**
- **Can be added to pre-calculated lighting values**
- **Subtractive lights *are* supported**

- **If you use 8 simultaneous lights then you may be doing too much work…**

NVIDIA

How fast is the lighting?

- Ambient is surely always free
- First light is usually ~free
- Ambient < Directional < Point < Spot

- What's this I hear about LOCALVIEWER?
  - Quality vs. speed
  - When should I prefer quality?

The DX7 Attenuation Model

Inspection should persuade you that the intensity doesn't fall to zero at any finite range. In DX7 you supply the three parameters a, b and c:

$$I = \frac{1}{a + bD + cD^2}$$

N.B. There are two errors in the DX7 docs which might lead you to believe otherwise. Firstly there is a claim that D is normalized (it isn't) and secondly there is a claim that the intensity typically varies from 1 to 0 across the light's range (it doesn't).

NVIDIA

Summary

- Use the D3D pipeline
- Tuning for hardware is straightforward
- The behavior is different in several ways
- DX7 is not exactly the same as DX6
- Correct use of VBs is critical

- Tuning for H/W is easier than for S/W

NVIDIA

Questions…

Richard Huddy

RichardH@nvidia.com

NVIDIA