Johns Hopkins
Engineering for Professionals

605.767 Applied Computer Graphics

Brian Russin

Module 11F
GPU Architecture

Graphics System Architecture

- Graphics systems architecture is a specialized branch of computer architecture
  - Many of the same efficiency techniques can be used
    - Pipelining and parallelism
      - Combinations of these are often used
    - Tradeoff between memory and computation
- Graphics computations required generally exceed capacity of single CPU
  - Even for modest complexity graphical scenes and frame rates
  - Note that some CPUs have support for geometry acceleration
    - Single-instruction Multiple Data (SIMD) allows efficient matrix and vector operations
- **Driver** is the interface between the application (in CPU) and graphics hardware
  - Example: OpenGL drivers for NVIDIA graphics card
  - OpenGL drivers must conform to OpenGL specification

History

- Evolution of PC graphics hardware has started from the end of the pipeline
  - Early graphics accelerators supported fast rasterization/span drawing
    - Draw interpolated or textured spans
      - Most performance to gain was realizable in the rasterization stage
  - Triangle setup added in 1996 – 3Dfx Voodoo 1
  - Geometry stage placed in hardware later
    - NVidia GeForce256 (1999) introduced hardware **Transform and Lighting (T&L)**
    - Note: many workstations already had advanced graphics hardware support
      - Silicon Graphics, HP, Evans and Sutherland
- Early 2000s, vertex and pixel shaders introduced the ability to program functionality within graphics hardware
  - Rather than a **fixed-function** pipeline
  - Modern graphics accelerator is often called a **Graphics Processing Unit** or **GPU**

Peripheral Display Processors

- Early trend in computer graphics was a separate graphics computer
  - Mainframe would send data to the graphics device
  - Ramtek, Evans and Sutherland were two of the more capable devices
    - Last Starfighter -nearly half the movie was computer graphics generated with a Ramtek driven by a Cray supercomputer
- Many early systems were called **graphics terminals**
  - Limited graphics capabilities by today's standards
  - Useful for plotting data
- Most of these were display list systems
- Display commands were loaded and stored in graphics device
  - Slowness of downloading commands and data to the graphics device
  - Scan conversion (polygon fill and line generation) supported
  - 3D effects generally in CPU

Scan Assist Architecture

- Common architecture in 2D workstations and basic PCs of 1990s
- General purpose CPU and a special purpose display processor
  - Pixel engine
- Pixel engine supports rasterization and scan conversion tasks
  - Bresenham line drawing
  - Shaded line support (for Gouraud shading acceleration)
  - Depth test logic and z-buffer
  - Scissoring (prevent pixel writes outside region)
  - Direct memory access for pixel transfers
  - Wide pixel bus