David Skilton: 2006

TECH2003 - Technology Fundamentals

Technology Case Study

Computer Game Graphics Technology

1.0 Introduction

The computer games industry is a multi-billion dollar market, with both hardware and software manufacturers fighting for market dominance. With high end graphics cards approaching prices upwards of £400, industry leaders in GPU[1] technology battle to release ever more powerful accelerator cards to eager high-end gamers. At the same time, software developers are creating new and fascinating graphical techniques, providing more realistic images in games, and pushing even the newest cards to the limit of their processing power.

We will be examining the ways that modern computer graphics hardware technology enables producers to develop immersive near-photo-realistic game worlds, and the software techniques involved in creating them. The range of hardware available, as well as the graphics techniques and effects used in high-end real-time 3d rendering are many and varied, but in this study we will examine some of the most frequently used and most effective.

2.0 Early development of computer game graphics

Early computer and video games featured 2D sprites of limited resolution, colour and animation. Processing power and memory were very limited, restricting the visuals to simple iconic representations of characters and game objects. As computing capabilities increased, as did the possibility for using 3D techniques for the creation of game environments, and the establishment of new genres such as the FPS (First-Person Shooter).

2.1 Early attempts at 3D graphics

Generally accepted as the world's first truly 3D game, Atari's *Battlezone* (1980) used simple wire-frame vector graphics to simulate a first-person perspective. The possibility for a first-person viewpoint was explorer further in Malcolm Evans' *3D Monster Maze* (1981), using 2D scaling pixels to create a similar effect. It was not until the huge popularity of id Software's W*olfenstein 3D* (1992) and later, *Doom* (1993) that the First-Person Shooter was recognised as an established genre in the world of computer and video games. Both games used a pseudo-3D effect with 3D scaling sprite characters, and signalled a huge change in the way that games were developed. id Software later went on to create *Quake* (1996), an entirely 3D game, comprised of polygonal and textured sprites and game world.

2.2 Early graphics hardware

At this time, the computer's CPU was used to handle the rendering of in-game graphics, but as 3D techniques became more advanced, the need for a separate dedicated graphics processor became apparent. The newly formed 3DFX company released the first gaming-

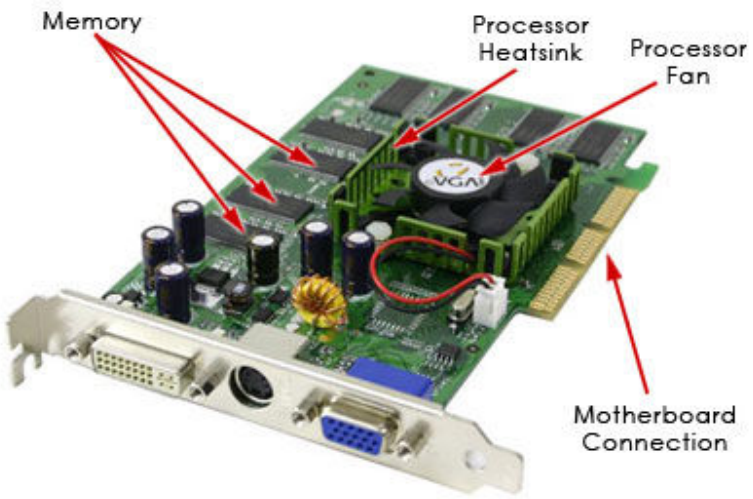
[1] Graphics Processing Unit

orientated graphics accelerator chipset, known as *Voodoo*. Plugged into the motherboard using the standard PCI interface, the graphics accelerator would activate when 3D rendering was required.

For effective communication between game programming, device drivers, and hardware, a standard API[2] was required. Voodoo cards at the time used 3DFX's *Glide* API, which was relatively intuitive to program for, resulting in market saturation throughout the 90's. Today, Microsoft's DirectX API is the most commonly (usually exclusively) used for Windows-based games. It has become the standard 3D development tool, for game designers, and is updated often as new graphical techniques are conceived (see sections 4.2 to 4.5).

3.0 Graphics Processing Units

The sole function of the GPU is to handle the rendering of graphics, leaving the main CPU to handle other tasks such as game workings and physics. Diagram 1 gives an example of a typical graphics card found in a gaming-spec PC.



graphics-card-5.jpg – http://computer.howstuffworks.com/graphics-card1.htm - (accessed 01/04/06)

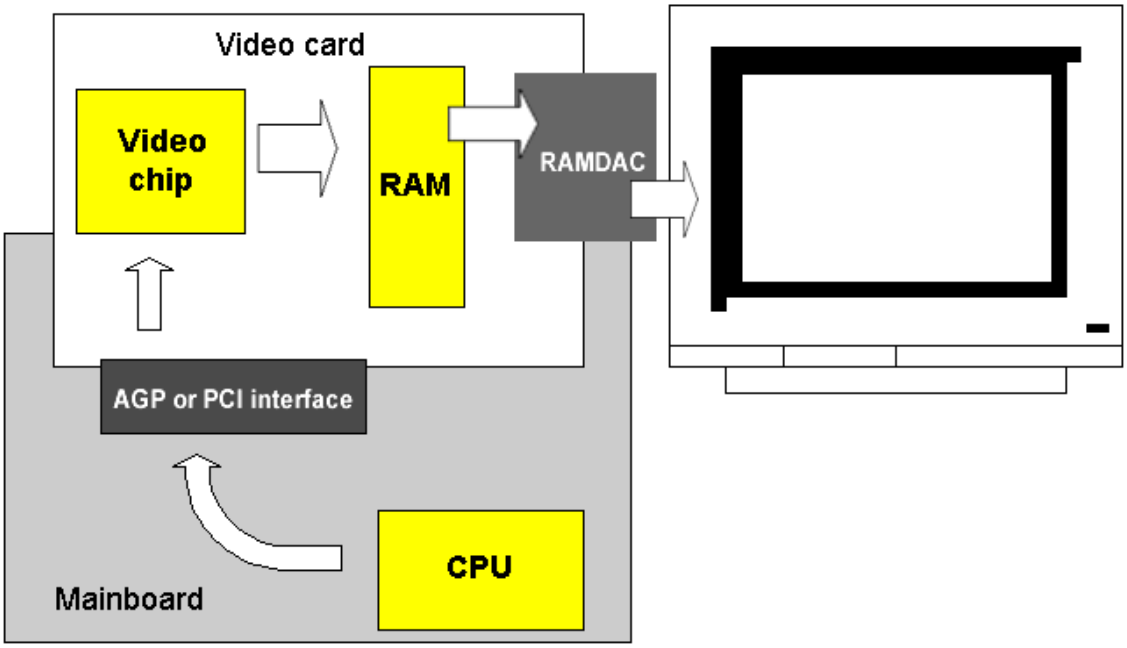
The main components are the VRAM (also known as frame buffer memory), Graphics Processor Unit with cooling equipment, and motherboard interface, which will use either the AGP standard, or the newer and faster PCI-Express[3] specification.

[2] Application Programming Interface – a library application for programs to request data and instructions.

[3] Originally the standard PCI (Peripheral Component Interconnect) interface was used, but bandwidth restrictions lead to the development of AGP (Accelerated Graphics Port – up to 8x bandwidth capability) and the most recent, PCI-Express (up to 16x).

3.1 Signal flow

When an application creates an image, the card's driver software acts as a go-between for the Operating System and graphics board, converting the data to a format that can be understood by the board's architecture. The driver then sends this formatted data into the board for rendering via the motherboard interface. The data is temporarily stored in the VRAM frame buffer then sent to be rendered by the GPU. The RAM digital-analogue converter then reformats the information into an analogue signal which can be sent to the monitor as an image.



U1929us.gif - http://www.karbosguide.com/hardware/module7b1.htm - Karbos Guide - The Video Card (accessed 01/04/06)

This process occurs around 20-90 times a second, depending on the graphical intensity of the render.

3.2 Graphics card specifications and benchmarking

The two GPU manufacturers prominent in the industry today are NVIDIA, with their *GeForce* chipset, and ATI, developers of the *Radeon* chipset. 3DFX and other companies such as Matrox established the market in the late 90s, but ATI and NVIDIA gained market dominance over the past 5 years. The two companies do not produce graphics cards, but the GPU systems that they utilise. Companies such as XFX, BFG, Leadtek, and MSI, among many others design and manufacture the actual cards, with varying, usually minimal differences between brands. These differences can include, but are not limited to:

- Varying levels of customer service and warranty,
- Free games software,

- Pre-set overclocked[4] settings,
- Different types of cooling systems.

Chipset specifications are noted in terms of core clock speed (the speed of the main GPU in MHz), pixel pipelines (components that transfer pixel information), memory capacity (VRAM – memory used to store image and texture data) and memory clock speed (memory speed measured in MHz).

The performance of graphics cards is measured in two ways. Firstly, numerical data of the number of pixels that can be rendered per second (fill rate, measured in MTexels/s[5]) is measured. This can be seen as theoretical performance, under conditions that may not reflect general use in games. The second measurement method is practical benchmarking, in which an identical recording of a game sequence is used to measure the average frames-per-second that the card is capable of producing. This method proves more valuable for buyers, as it provides a more realistic indication of performance in a gaming environment, and is the standard procedure for hardware reviewers performing comparison tests (found in PC magazines and websites).

The following table shows specifications and benchmarking performance of various NVIDIA chipsets over the past 9 years, giving a rough indication of the rate of technological increase.

Year	Model	Interface	Core clock	Pixel Pipelines	Max Memory	Memory clock	MTexels/s
1997	Riva 128	AGP 1x	100MHz	1	4MB	100MHz	100
2000	GeForce	AGP 4x	120MHz	4	128MB	300MHz	480
2003	GeForce FX	AGP 8x	250MHz	2	256MB	400MHz	1000
2006	GeForce 7	PCI-E 16x	450MHz	4	256MB	666MHz	1800

Data taken from: *Comparison of NVIDIA Graphics Processing Units*
http://en.wikipedia.org/wiki/comparison_of_NVIDIA_Graphics_Processing_Units (accessed 01/04/06)

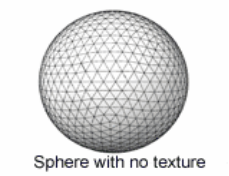
4.0 Modern Graphics Techniques

Microsoft's DirectX has become the standard API that developers use for coding graphics for Windows-based PCs. Using the standard Direct3D tools within the DirectX API, graphics have evolved greatly over the past 10 years, each new graphical feature improving an aspect of realism in game worlds. The following is a brief rundown of the key features introduced with each iteration of DirectX.

[4] The process of increasing a processor's clockspeed by increasing input voltage. This increases heat, which must be handled carefully to avoid damage to core components.
[5] Multi-textured pixels.

4.1 Standard techniques – Texturing and Antialiasing

At the fundamental level, three-dimensional computer graphics are comprised of rendered polygons on a 2D plane, simulating depth using perspective. Every part of the game world is a 3D polygon model with bitmap images mapped onto the surfaces. This is known as texture mapping.

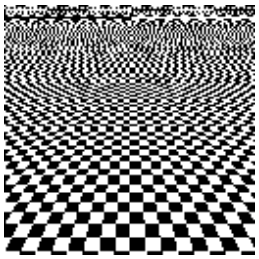

Sphere with no texture

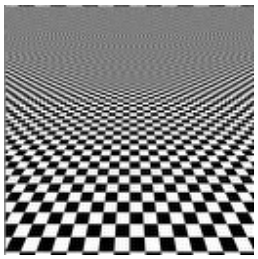

Texture image


Sphere with texture

TextureMapping.png – http://wikipedia.org/wiki/Texture_mapping - accessed (01/04/06)

One of the first image enhancing techniques developed after basic textured polygon graphics was Anti-aliasing. Anti-aliasing techniques work to enhance the image in two ways. Firstly, it reduces the jagged edges[6] produced by visible pixels in a mid-resolution image (see image 1.), and secondly, it can recreate visible elements that are invisible due to being smaller than the onscreen pixel size. This is achieved by sampling the image multiple times[7] then down-sampling back to screen resolution using interpolation, which creates mid-range values for in-between pixels, increasing image smoothness and providing finer detail (see image 2.) Of course, this enhancement is at a cost of frame-rate performance.

1. 

2. 

http://en.wikipedia.org/wiki/antialiasing (accessed 01/04/06)

4.2 DirectX 6.0 – Multitexturing and Hardware T&L

With the introduction of Direct3D programming using the DirectX API, came an advanced texturing ability known as Multitexturing. This technique involves the graphics card applying more than one texture per render pass. Multiple texture layers are applied to an object without having to load the object multiple times. This method allows the rendering of complex materials with minimal loss of speed (compared to multi-pass rendering techniques).

[6] Also known as 'stepping'. Most visible on diagonal lines.

[7] Usually x2 (twice the screen resolution) or x4 (quadruple the screen resolution.) Up to x8 on some high-end graphics cards.

Hardware Transform and Lighting allows for an accurate conversion of the spatial coordinates of a 3D space on a 2D plane (output image) and can calculate advanced lighting on objects in a scene, calculating the resulting colour of surrounding objects.

4.4 DirectX 8.0/8.1 – Shader model 1.1, Pixel Shaders and Vertex Shaders

The arrival of DirectX 8.0 introduced Shaders to the standard. Pixel and Vertex Shaders determine the final surface properties of an image or object, attempting to recreate the various effects of reflection, refraction, shadowing and displacement found in the real world. Full screen pixel shaders create depth-of-focus, motion blur and heat haze effects.

Atisushi01.jpg – www.tommti-systems.com (accessed 01/04/06)

4.5 DirectX 9.0/c – Shader Model 2.0 and 3.0

The most recent version of DirectX features upgraded and refined Shader methods. Shader Model 3.0 introduced HDR (High Dynamic Range) rendering. This technique alters the way that light is displayed on-screen, allowing light areas to clip into white, simulating areas that are too bright for the human eye to see. Dynamic changes are also made to the amount of light in a given area, simulating the eye's adjustment to said light levels.

1)  2) 

Farcrynohdr.jpg and Farcryhdr.jpg – http://en.wikipedia.org/wiki/high_dynamic_range_rendering (accessed 01/04/06)

Image 1 demonstrates the traditional effects of lighting in a scene, where image 2 features HDR rendering. A greater, more accurate contrast is achieved, with white areas clipping and blooming where the human eye would fail to adjust in such a darkened environment.

5.0 Current technological limits

The main limit of graphics technology, and indeed with processing in general, is clock speed. As processors are manufactured with ever more intricate procedures, faster chips are able to be developed. However, as speed increases, as does heat output. It is becoming increasingly difficult to maintain safe working temperatures[8], and card manufacturers must take care to supply adequate cooling systems. High-end users who wish to overclock their core and memory speeds face even greater problems, and often require the advanced cooling provided by refrigerated water systems.

Power consumption is also an issue. A modern graphics card can be the single biggest drain of power in a computer system. Some high level cards require a power supply unit with an output power of over 400W in order to run effectively alongside hard drives and other devices in a PC system.

6.0 New and Future Developments

The current technology available to high-end users is SLI[9]/Crossfire (NVIDIA and ATI's version respectively). This system allows cards to be run in tandem on capable boards (usually with two or more PCI-E slots), each card rendering a separate half of the screen. SLI and Crossfire is not currently supported by many games however, and purchasing multiple high-end graphics cards can be expensive, so the method is generally only used by high-end users. Currently in development are quadruple-SLI systems, allowing 4 identical graphics cards to be used simultaneously for maximum performance.

DirectX 10 will signify the arrival of the Windows operating system to the 3D realm. Using the same technology as games do, the *Windows Vista* (Microsoft's upcoming operating system) user interface environment will be entirely rendered by the system's dedicated graphics card. This provides a huge shift

[8] GPUs can run at upwards of 80 degrees Celsius at full load.
[9] Synchronised Link Interface.

in the ways that 3D is incorporated in a system; with DirectX 10 moving away from the driver-based architecture of previous versions and becoming an integral part of the operating system itself.

In the same way that graphics processing switched to dedicated hardware, there is soon to be a change in the way that in-game physics are handled. With games becoming increasingly reliant on heavy physics calculations to determine animation and object properties, separate physics processing units (PPUs) will eventually become standard hardware, either as a standalone dedicated card, or a separate chip incorporated on the motherboard or graphics card.

7.0 Conclusion

The techniques used for the creation of environments and objects in games software has advanced greatly over the past decade. Processing power has multiplied many times over, and we are starting to experience graphics with near-photo-realistic quality, allowing developers to create truly immersive game environments. Although graphics processing hardware is a relatively new technology, huge advancements have been, and continue to be made. New and exciting prospects await games as well as developers, as more advanced techniques emulating real world properties are developed.

Although high-end graphics capabilities are costly and only for true enthusiasts, the future integration of operating systems and graphics hardware may narrow the gap. Such implementation, as well as the continually growing popularity of computer and video games will certainly ensure that future advancements continue to be made, and that the industry remains strong.

Bibliography

Wikipedia - http://www.wikipedia.org - Various sections: (accessed 01/04/06)

3DFX

Antialiasing

Comparison of ATI Graphics Processing Units

DirectX

Full screen effect

Graphics card

Graphics processing unit

High Dynamic Range Rendering

List of computer graphics and descriptive geometry topics

Shader

Texture mapping

Transform and lighting

How Stuff Works - http://computer.howstuffworks.com/graphics-card1.htm (accessed 01/04/06)

KarbosGuide - http://www.karbosguide.com/hardware/module7b1.htm (accessed 01/04/06)

Acme How-To - http://www.acmehowto.com/howto/pc/hardware/graphics/overview.php (accessed 02/04/06)

PC World - http://www.pcworld.com/howto/article/0,aid,31050,pg,1,00.asp (accessed 01/04/06)

AnandTech - http://www.anandtech.com/showdoc.html?i=1288 (accessed 01/04/06)

My PC Plus - http://www.mycplus.com/c.asp?ID=9 (accessed 01/04/06)

Bit-Tech - http://www.bit-tech.net/hardware/2005/07/04/aliasing_filtering/3.html (accessed 02/04/06)

References

Atisushi01.jpg – www.tommti-systems.com (accessed 01/04/06)

Comparison of NVIDIA Graphics Processing Units -

http://en.wikipedia.org/wiki/comparison_of_NVIDIA_Graphics_Processing_Units (accessed 01/04/06)

graphics-card-5.jpg – http://computer.howstuffworks.com/graphics-card1.htm - (accessed 01/04/06)

Farcrynohdr.jpg and Farcryhdr.jpg – http://en.wikipedia.org/wiki/high_dynamic_range_rendering (accessed 01/04/06)

TextureMapping.png – http://wikipedia.org/wiki/Texture_mapping - accessed (01/04/06)

U1929us.gif - http://www.karbosguide.com/hardware/module7b1.htm - Karbos Guide - The Video Card (accessed 01/04/06)