MRU Tech Notes
Rates of frames

Outpost, 8 Dicembre 2007

Marco Ruocco

Preamble

Every bottle has a bottleneck, and the same is true for computer systems in their graphical subsystems, which are the focus of this MRU technical note. The following is a qualitative discussion, or rather a conceptual map, about the technical constraints which usually bog down the gaming experience of hard core PC gamers as well as the Christmas-shopping choices of the slightly less experts on the matter.

Here we are talking video cards, graphic cards, graphical processing units (GPU), digital buses, digital, analog, output, refresh rate. This is a 100% technical article in concept. For a reference mark on these terms, look up the Appendix Zero at the bentonic bottom of this article.

Introduction

All computer games using 3D graphics make the computer calculate and output screen images at blazing-fast speeds: the amount of speed is measured in frames per second (FPS), that is, the number of times the screen is updated or refreshed in order to give the user the illusion of motion and change in the virtual environment in which she is playing. As a rule-of-thumb, the higher the framerate, the better the gaming experience is - although this is not always true.

New video cards deployed on the Western markets by Asian manifacturers approximately double the maximum framerate possible on our western PCs every year or so. However, new games are constantly released and they are more and more complex as the time passes, requiring more hardware resources, then decreasing the actual framerate from the original top value of the time of purchase. Every new game sets a new equilibrium point between the capability of the hardware and the functioning requirements of software running on that hardware.

Bottlenecks

Here we illustrate how the power of a videocard is scaled down as the video signal metaphorically flows downstream from the GPU (Graphical Processing Unit) to the visual river mouth (of the PC screen) and it is finally seen by the ocean's viewer (the gamer).

1) The screen frames produced by the Graphical Processing Unit (GPU) are sent to the video RAM at a blazing-fast speed. Speaking of bottlenecks, here the neck of the bottle is as wide as the bottom of the bottle – there is no data loss occurring, and this freeway has no speed limit (here at least).
2) From the video memory on the graphic board, screenframes are passed to various circuits such as the RAMDACs (for analogue RGB output) and Silicon Image chipsets (for DVI outputs). The maximum bandwidth at this stage is dramatically reduced from Neck 1, but it remains high. The throughput consists in up to 60 screenframes per second, and each screenframe may be up to 2 Megapixels in size. This is a second Neck, as restrictive as a beautiful pearly necklace - that is, this is not a harmful thing for gaming.
3) The video signal is then transported to the PC maintaining the bandwidth unchanged, via DVI or RGB short cables. Data loss is only due to electrical interferences, but it is not important at all.

4) The video signal is then input into the decoder within the monitor, and passed onscreen as digital images, at the correct resolution and refresh rate, as set on the PC and on the monitor button options.
5) The biggest issue is what reaches the eye of the beholder in front of the screen, which, after all, is just a piece of metal and glass. The eye is not a a camera-like instrument, but a visual system relying on perceptions and adaptive head-and-body motions. Psychophysical research has demonstrated that the human eye, in optimal conditions, can perceive refresh rates at no higher rates than 16 frames per second. In challenging conditions (such as, oblique angle of vision, specific complexity of environment, survival-related pressures, etc.) this measure may be higher – but no scientific research has been published on these topics *yet*.
6) Response time of monitors is expressed in milliseconds. This is a pure myth, probably due to wrong measurements at the source of the speed of LEDs to turn from black to white and back. This issue is a confound and will be dropped altogether.

Rationale

In case a future GeForce 9950 Ultra will output 255 frames per second in the game Doom3 at 1600x1200 resolution, those refresh feats are actually lost at Neck2, as the RAMDAC cannot physically cope with that influx of bytes. What gets to the user is the maximum throughput of the monitor screen, that is, the screen refresh rate (normally it is 60hz). Our visual system needs less than that, just about 24 fps in any condition of viewing (e.g. Cinema screens use films with 24 fps refresh rates, albeit with a light-emission based system).

Conclusive remarks

Our impression is that the most sought feature of any video adapter is steadyness in framerate, rather than "maximum framerate". A constant refresh of 24 frames per second would make any card the ideal card ever. Unfortunately, with the advent of Windows Vista, such reliability has collapsed for unknown reasons related to internal choices of implementation. Games would inevitabily drop from 200 fps to 1 fps in the most critical times of our gaming experiences, but this applies only to Windows Vista. Downgrading to XP is suggested and legal from a license standpoint.

Take-home message

MRU believes that once the quantitative revolution of pushing the envelope of maximum framerate and performance will be offloaded from the to-do list of video chipmakers, we all will be able to concentrate on visual quality, and look, device-strong, inside our digital windowpanes, onto what lies beyond, safely at one's own workbench.

Appendix 0

1) http://www.wikipedia.org
2) Ruocco, M. (2004). Human Psychological Response to Landscape Visual Filtering in animation design. MA Thesis, UC Santa Barbara, Geography Department. Chairs: Prof. Montello, Prof. Couclelis, Prof. Goodchild.