CS 529

Fundamentals of Game Development

1. RTIS?

1.1. Resolution

- What is a pixel?
  - A pixel is the smallest element of a picture.
  - In other words, it is the smallest unit of visual display that can be used to build an image.
  - The word pixel is derived from picture element.
  - Pixels are visible as tiny squares when an image is enlarged.
  - The higher the number of pixels in an image, the better the resolution.

- How is the resolution defined?
  - The number of pixels that the display card can support.
  - As the resolution increases, the quality of the picture increases, as well as the memory used to hold all the information concerning that particular picture.
  - Screen resolution refers to the number of pixels across and down delivered by the display card circuitry.
  - Nowadays, the display card circuitry known as the graphic card can support different resolutions.
  - A typical graphic card can handle resolutions from 640x480 to 1600x1200.

1.2. RGB concept

- How is a color composed?
  - All colors are made of three basic colors: red, green, and blue.
  - The intensity of each of the three basic colors creates another color.
  - Red = 0, Green = 0, and Blue = 0 will produce black.
  - Full red, full green and full blue will produce white.

- Color bit Mode
  - 8 bit (2^8 = 256 colors)
  - 16 bit (2^{16} = 65,536 colors)
  - 24 bit (2^{24} = 16,777,216 colors)

- 32 bit (2^{32} = 4,294,967,295 colors)

1.3. CRT concept

- What is the cathode-ray tube?
  - A vacuum tube in which a cathode (electron gun) emits electrons that are accelerated as a beam.

- How does it work?
  - Electrons strike the phosphor coating on the tube.
  - Light is emitted for a short period.
  - The direction of the beam is controlled by two pairs of deflection plates, or by a magnetic field generated by coils
  - The output of the computer is converted by a digital-to-analog converter.
  - The output of the computer is converted to voltages
  - The beam can move along the x and y directions.
  - The intensity of the beam can be turned off.
  - H-blank is the time needed to raster the next row.
  - V-blank is the time needed between the last pixel of the last row and the first pixel of the first row.
  - In order to see a steady image, the same path must be refreshed by the beam at least 50 times a second.
  - Colored CRTs have three different colored phosphors: red, green, and blue.
  - Consequently, there are three electron beams per pixel.

- The difference between interlaced and non-interlaced
  - The refresh cycle of the interlaced display monitor is divided into two frames.
  - All the odd-numbered scan lines are displayed in the first frame, and the remaining lines are displayed in the second frame.
  - When those two frames are drawn in succession, the mind creates an image of the combination of the frames.
  - The non-interlaced display monitor refreshes all horizontal lines sequentially.

1.4. Refresh Rate

- The number of times, per second, that images are refreshed on the display, measured in hertz (Hz).
- 60 Hertz is considered the minimum acceptable level.
- Do not confuse it with the frame rate of the application. They are totally different. The frame rate measures the speed a video source can provide an entire frame for the display.
- Even if the frame itself is not changing, the monitor keeps redrawing it over and over again
- Example:
  - The frame rate of a movie is usually 24, which means it has 24 different frames each second
  - But movie projectors usually have a 48 or 72 Hz refresh rate, which means they project the same frame twice or three times before projecting the next frame

1.5. Refresh Rate & Frame Rate

- As mentioned previously, the refresh rate of the monitor is totally different than the frame rate of an application
- The monitor and the video card do not have to be in sync, which means the refresh rate can be different from the application's frame rate
- Each time the monitor needs to refresh itself, it takes the content of the primary frame buffer (array of colors found on the video card) and displays it

1.6. Out of sync?

- What if the refresh rate of the monitor and the application FPS are out of sync? (And this is mostly the case).
- This means that while the application is producing a certain number of frames per second, the monitor is actually displaying a different number!
- If the application's FPS is different than the monitor's refresh rate, a tearing effect will take place.
- This occurs because when the monitor is refreshing itself, it ends up getting 2 or more overlapped frames.
- If the application's frame rate is way greater than the monitor's refresh rates, some frames will be completely missed, because newer frames got drawn on top of them in the frame buffer (graphics card) before the monitor got the chance to display them.
- Tearing may be somehow acceptable for some games, but it will be highly noticeable and disturbing in fast paced games, because the visual difference between two consecutive frames is relatively high.

1.7. Vertical sync

- Vertical sync (or simply vSync), the graphics card is told to synchronize itself with the monitor's refresh rate.
- This is done by making the graphics card *wait* for the monitor to signal that it is ready to display a new frame before actually supplying a new one.
- This way, the application can't produce frames more than the monitor can handle, therefore the content of the graphics card's frame buffer will always contain a whole frame, and not several overlapping frames, thus eliminating the tearing effect.
- Assuming the monitor's refresh rate is 60 Hz, any running application with vSync turned on will be capable of producing a maximum of 60 frames per second.
- So far vSync seems like a perfect solution, but what if the application's frame rate is less than the monitor's refresh rate and vSync is turned on?
- When the application generates a new frame, it would have missed the monitor's refresh time by a small amount of time, therefore it doesn't generate the next frame until the monitor displays the current frame.
- In other words, each 2 monitor refreshes will display the same image, which reduces the application's frame rate to half the monitor's refresh rate.
- Example:
  - Application FPS: 50. Frame generation time: 1/50 = 0.02 seconds
  - Monitor refresh rate: 60. Frame display time: 1/60 = 0.016 seconds
  - Let's see the state of the primary frame buffer (PFB) each time the monitor refreshes itself.

Time	Primary frame buffer	Monitor
0.0	Blank	Nothing yet
0.016 (Monitor refresh time)	Blank	Blank
0.02	Contains frame 1	Blank
0.032 (Monitor refresh time)	Contains frame 1	Displays frame 1 (Now the game can start working on frame 2)
0.048 (Monitor refresh time)	Contains frame 1	Displays frame 1
0.052	Contains frame 2	Still displaying frame 1
0.064 (Monitor refresh time)	Contains frame 2	Displays frame 2 (Now the game can start working on frame 3)
0.08 (Monitor refresh time)	Contains frame 2	Displays frame 2
0.084	Contains frame 3	Still displaying frame 2
0.096 (Monitor refresh time)	Contains frame 3	Displays frame 3 (Now the game can start working on frame 4)

- As you can see, the monitor refreshes itself twice using each frame, which ultimately drops the application's frame rate to 30 (Monitor refresh rate / 2)

- Following the same logic, if the application's frame rate is already lower than 30, it will be forced to drop to 20 (Monitor refresh rate / 3), and so on..

1.2. LCD monitors

- Unlike CRT monitors, LCD monitors do not have an electron gun
- It contains an array of liquid crystals, between two layers of polarized glass
- Light is sent from the back of the first glass layer passing through the crystals
- On the other hand, the crystals' orientations are modified in order to alter the light that passes through them to the screen
- Therefore, an LCD screen doesn't need to refresh the entire display all the time, it can only modify a part of it by changing the shape of the correspondent crystals
- The time needed by the LCD monitor to change the shape of the crystals is called "Response time". The response time (in milliseconds) is determined by calculating the amount of time needed by the LCD to go from full black to full white
- Logically, the lower the response time, the better. A lower response time means that the LCD monitor can react faster to color changes coming from the application
- For compatibility reasons with games and hardware, a refresh rate can be set for LCD monitors, especially for games where vSync is enabled
- Remember that when vSync is enabled, the game FPS is locked to the monitor's refresh rate. This is why LCD monitors need to emulate the CRT's refresh rate, although they don't have this limitation

