Notes on Latency

[WIP]

For the purposes of this writeup, we'll consider latency to be the delta or time between a person's action and the response or display from the software that captures their image.

Most displays run at 60 frames per second, which means they are showing an image every (approximately) 16 milliseconds. When a camera is capturing an image, delay is introduced at several stages - the first and largest of importance is the camera's exposure time. Cameras need to leave their shutter open for a period of time to collect light on the sensor - the longer the shutter is open, the more light is collected. At 60fps, the ideal is a shutter speed of 1/60s, but shutter speeds can be longer to allow more light in for a darker environment. Additionally, latency can then be introduced by a camera's own buffering and processing, transmission to computer and subsequent writing to computer memory, and finally your own software's processing.

Some cameras that advertise "low latency" are at around 35ms, which means in a best case scenario, you're seeing your reaction almost 2 frames after you've performed an action.

Latency is especially noticable for audio applications and less so for visual. Musicians begin to have trouble with latency that is larger than 8-12ms. If you're planning on using a camera as a trigger for something like an interactive drumming station, you may find that the latency is unacceptable for certain kinds of playing.

Here is a great Reddit thread on computer vision latency

Revision #1
Created 2025-05-07 19:21:05 UTC by Blair Neal
Updated 2025-05-07 19:21:41 UTC by Blair Neal