Interpolation

Sunday, October 05, 2025 7:55 PM

Currently this idea entails taking the 30fps video and interpolating to a higher fps and then down sampling back to 30fps using these new frames.

What does sampling phase mean?

Let's say the original video is 30fps (F_0) -> you interpolate to 60fps (F_u). When you downsample by **m** = (F_0/F_u)

If you keep every m-th frame at the index p = {0,…,m-1}, the kept frames occur at the times:
$T(j) = (p+m*j)/F_u$

This would be exactly the same as the original cadence but with a constant time offset of:
Effective time shift = p / F_u (seconds)

The duration won't change but the entire video timeline is shifted relative to the original which can cause issue when paired with audio

Example

Going from 30 to 60 -> m=2

p=0 → Δt = 0 ms

p=1 → Δt = 1/60 s ≈ **16.67 ms** forward shift (video leads audio if you don’t adjust audio)

What are my options

To get the audio back on track exactly the audio could be shifted the same amount

I could take inspiration on Naomi Hartes paper where (start/middle/end) of an utterance was masked one at a time but instead the audio is shifted slightly forward or slightly backwards

What metric can be used to assess the interpolation (upsampling quality)

It was said that interpolation tools can be provided within the college but whats the best metric to judge the quality?

Assess the tool

I could down sample the original video to 15fps and ask it to upsample to 30 and then run a comparison to the original 30 fps video