Slow Motion Video

MediaTek Technology White Paper

May 2015

Introducing Slow Motion Video

MediaTek's Slow Motion Video technology allows the device user to easily capture, replay, and share a fast moving object any time. Slow Motion Video can record video at a high frame rate such as 120fps, 180fps or higher, playback in slower speed such as 1/2x,1/4x, *and 1/16x*. And, the user can easily share this slow motion video.

High frame rate recording allow users to capture more details and wonderful moments for a fast moving object. Compared with normal camera recording, which only can capture 24fps – 30fps, high frame-rate recording can capture as much as 120,180fps, which allows for more fluent recording.

Mediatek provides a whole solution for high frame-rate recording, so that business partners can easily combine their high frame rate camera sensors with Mediatek ICs to provide many new wonderful features. Now VGA@120fps for MT6752, HD@180fps for MT6595 and MT6795 are supported, and higher, such as HD@240fps, will be available on later MediaTek octa-core ICs.

Slow Motion Video playback allows user to slow down the playback speed so as to enjoy a fully playback experience. Playing the high frame rate recorded video will bring much more fluent watching experience. But as we all know, the human video processing capability is limited, so humans cannot capture details of a fast moving object if replayed in 1x speed. By using the slow motion playback, the user can not only distinguish each amazing moment but also can enjoy interesting viewing experience.

Now Mediatek can support fluent slow motion playback with slowest speed of 1/16x. What's more, Mediatek Slow Motion Video not only supports slowing down the playback speed but can also adjust the effect interval. The user can just slow down the interested interval while playing other intervals at normal speed.

The **Slow Motion Video Share** feature allows the user to create the slow motion effect on video and *share* the effected video to friends. The Sharing feature allows distributing the interesting video by social media like Youtube, Facebook, email, or radio transmission technology such as Bluetooth.

With Mediatek Slow Motion Share function the user can share exactly what she got during playback to friends. The user can adjust the slow motion effect section and set the speed before sharing.

Not all devices can support high frame-rate video playback fluently. To ensure that all sharing friends using different terminals can enjoy the same user-created experience, and to increase the sharing speed, the Mediatek Slow Motion Share function will downgrade the video with slow motion effect to normal frame rate of 30fps before sharing.

Technology Problems

Implementing all the above Slow Motion Video functions presents many technology challenges, as listed here below.

High Frame Rate Recording

High frame rate recording not only requires powerful hardware, but also needs little SW overhead to raise the HW utilization, which requires many optimizations and tune efforts

For video encoding, the HW usage could be divided into 3 parts: Image Signal Processor (ISP), Multimedia Data Path (MDP), and Video Encoder (VENC). The HW block diagram is shown in the following Figure 1.

Figure 1. Video Recording HW Block Diagram

Requiring high SW overhead in each HW stage will obviously bring disaster for high frame rate recording. To achieve high frame rate recording, the first and the most important technology problem is the SW overhead penalty for each frame.

Traditional video HW encoder architecture imposes a pre-processing time and post-processing time SW overhead penalty for each frame.

- Pre-Processing: before triggering HW encoding a frame, the encoder driver needs to provide useful information to HW (e.g. I frame period, CBR/VBR …).
- Post-Processing: after HW encoding, useful information such as key frame information, encoded bi-stream byte count is stored in the file writer.
- According to our experience (MT6795), we need at least ***1.5ms*** for pre-processing and post-processing for each frame.

In normal frame rate recording, the SW overhead is insignificant, as shown below. e.g. 4K2K@30fps, SW/HW = 1.5/31.8 = 4.7%.

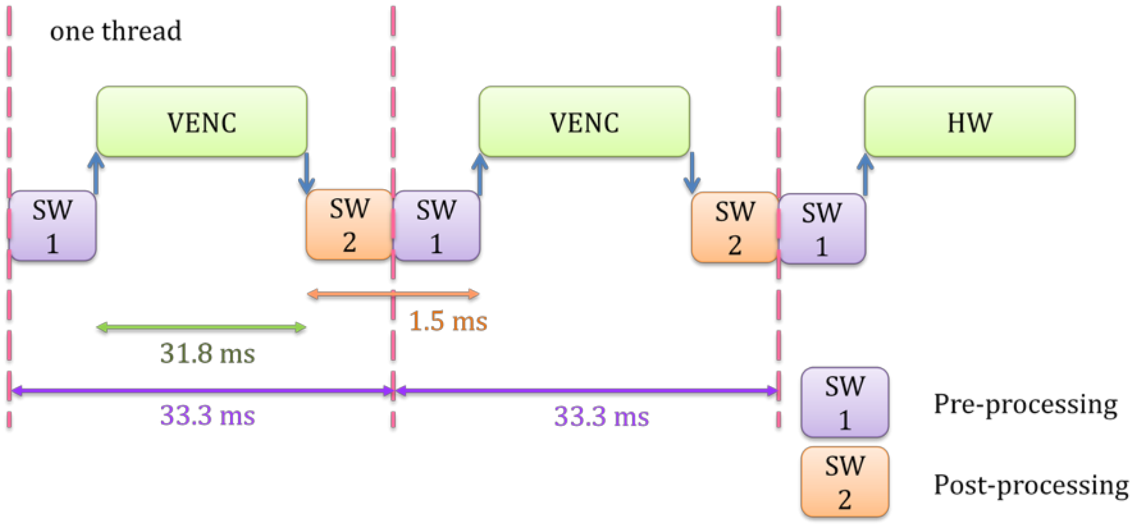

Figure 2. SW Overhead Penalty for Each Frame of 30fps Recording

But in high frame rate recording, the SW overhead is obvious, as shown in Figure 3. e.g. HD@180fps, SW/HW = 1.5/4 = 37.5%. The SW overhead will degrade the HW utilization rate and become the main bottleneck to achieve higher than HD@180fps. Minimizing SW overhead is very important and challenging.

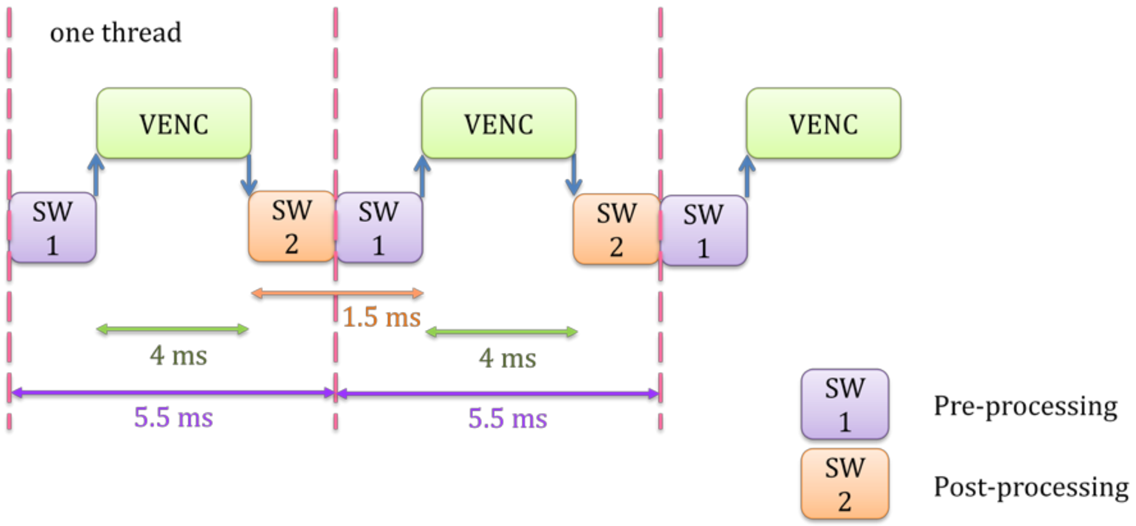

Figure 3. SW Overhead Penalty for Each Frame of 180fps Recording

Slow Motion Video Playback

Slow Motion Video Playback requires slowing down the video track *and* the audio track. The problem is how to maintain the audio quality, without changing the timbre.

To change the audio playback speed, the audio duration needs to be extended or shortened. The simplest way to modify audio duration is to resample it. Resample involves rebuilding a continuous waveform from its original samples at different rate. Thus, new samples will be added if an audio is to be extended, and vice versa, samples will be cut if an audio is to be shortened. When resampled audio is played at the original sample rate, it sounds faster or slower.

However slowing down the resampled audio will lower the pitch, and speeding up it will raise the pitch. Resampling is a simple way to modify audio duration, but real audio information is lost.

The MediaTek Solution & Algorithm

A series of optimizations and algorithms are adopted by Mediatek for the Slow Motion Video Feature.

High Frame Rate Recording Solution

To fully utilize HW capability, Mediatek has applied series of optimizations on the video recording path, including VENC and MDP. Finally the SW overhead of the recording path is minimized to nearly 0.

To fully utilize the VENC HW, Mediatek pipelines the SW and HW between MDP and VENC, shown as Figure 4. The VENC and MDP could be pipelined via Command Queue. This design raises the HW utilization and reduces VENC SW overhead. With this Technology the VENC HW register can be configured when the MDP HW is working.

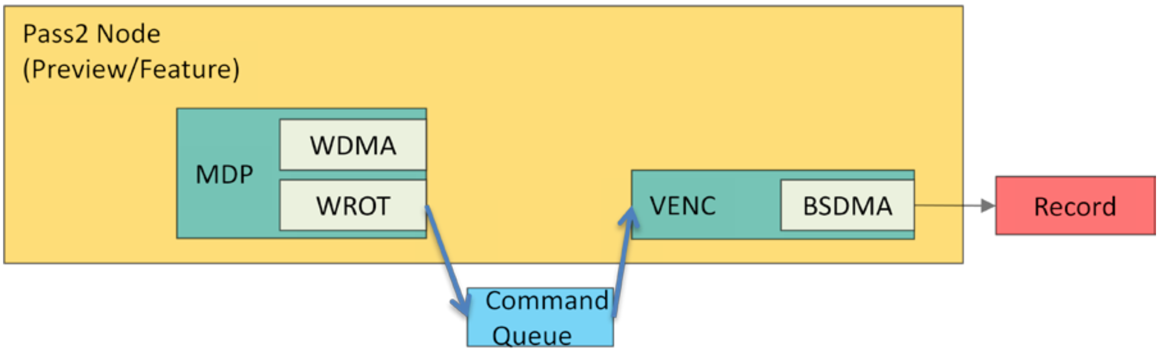

Figure 4. New Pipeline Design between MDP and VENC

Figure 5 and Figure 6 show the different results between before and after applying “SW/HW pipeline between MDP and VENC”.

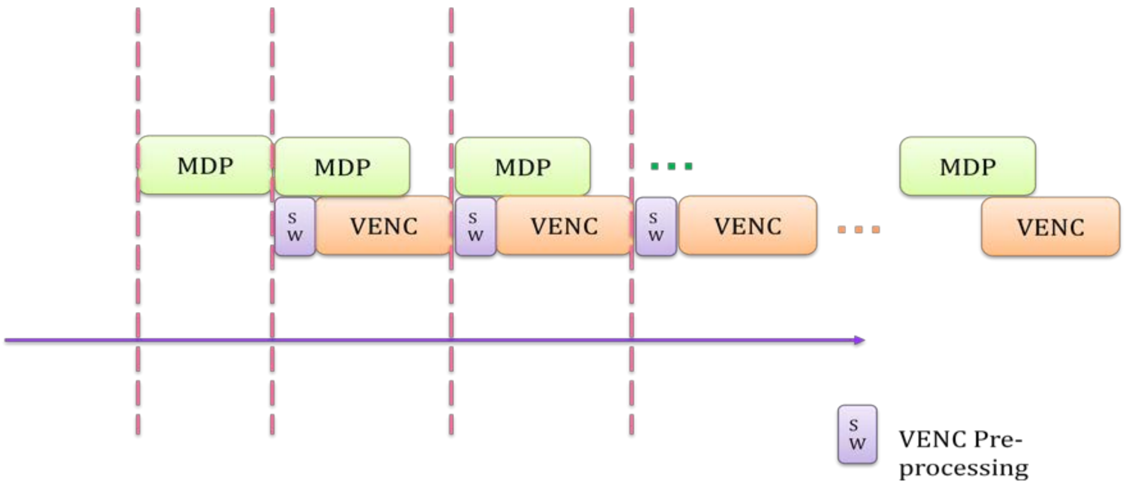

Figure 5. VENC without Using Command Queue

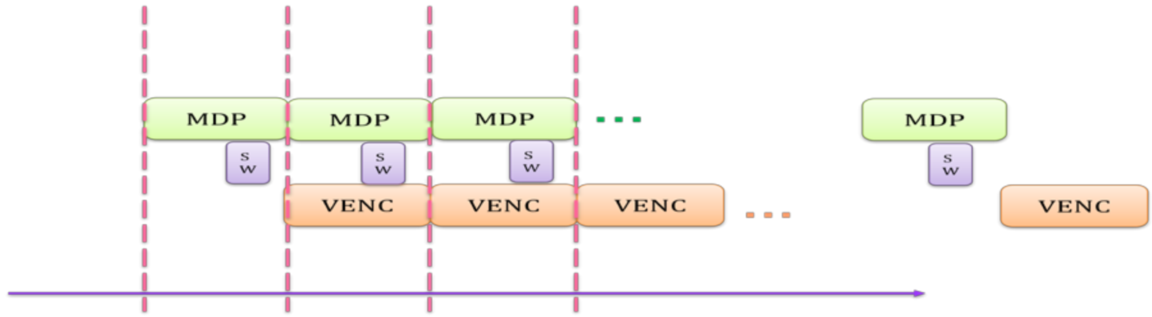

Figure 6. SW/HW Pipeline between MDP and VENC

In Figure 5 VENC has a SW overhead penalty after MDP for each frame. Figure 6 shows, on the other hand, that VENC can start immediately after MDP processing with adopting "pipeline between MDP and VENC".

What is more, in MediaTek's video HW encoder architecture, the encoding task is split into three threads: HW thread, SW1 thread, SW2 thread. In order to hide SW overhead into HW processing time, SW1 and SW2 phases need to be parallel with HW processing, as shown in Figure 7 below:

- Prepare HW needed information of next frames early
- Collect HW output information even if HW is processing

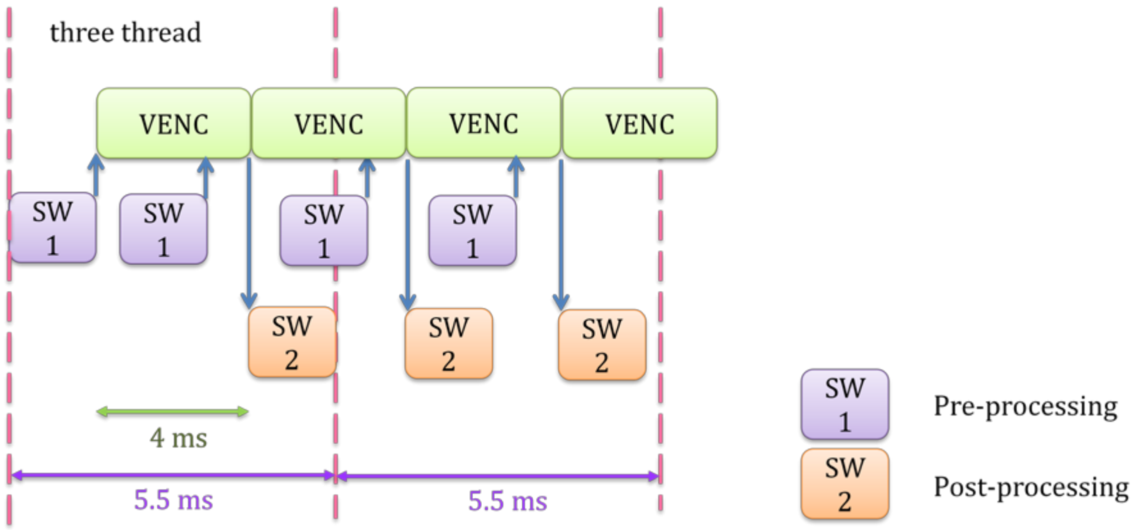

Figure 7. SW Overhead & HW Processing Parallel Solution

As a result, the VENC SW overhead can be hidden, achieving the goal with of 0 SW overhead.

To fully utilize the MDP HW, asynchronous MDP interface is introduced, as shown below in Figure 8.

- To increase hardware throughput, the MDP driver applies a new design for software and hardware pipeline optimization.
- The new design allows MDP SW driver configuration to perform asynchronously for every frame. In other words, the SW driver could configure the next frame while HW still operates the last frame.
- It reduces the user waiting time, and increases the HW utilization.

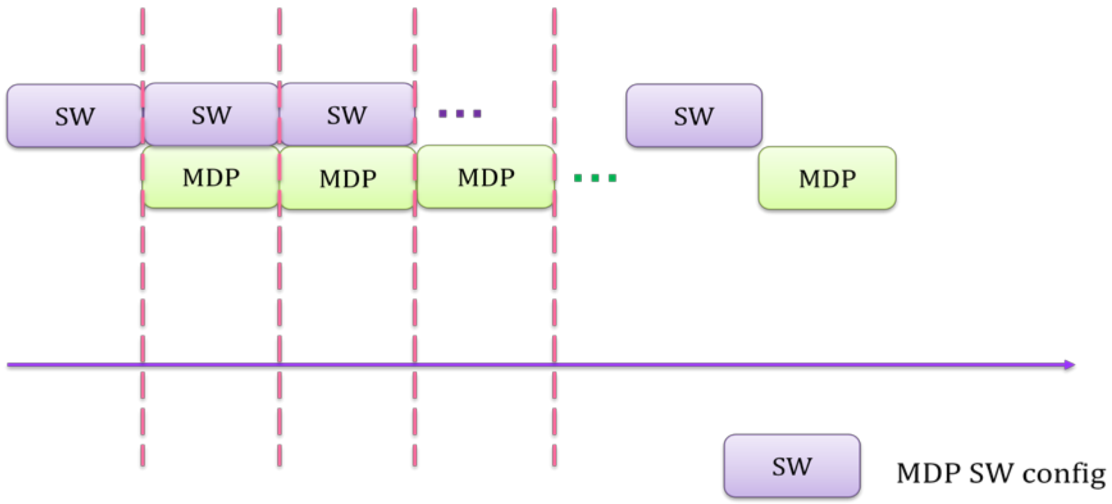

Figure 8. Asynchronous MDP Interface

Audio Time Stretch Algorithm

To maintain the audio quality while slowing down the audio playback speed, Mediatek has adopted an Audio Time Stretch algorithm instead of simple resampling.

Audio Time Stretch technology can retain pitch while prolonging or shortening audio duration. Among the Audio Time Stretch algorithms, Mediatek's exploits the well known time stretch algorithm Pitch Synchronous-OverLap-Add (PSOLA) for its high efficiency and simplicity, and improves PSOLA algorithm with our enhancement.

The PSOLA algorithm captures audio waveform in time domain. It tries to find a wave form that has the same pitch and overlap, and adds it to prolong or shorten the duration. The whole process can be divided into Pitch Detection and Cross Fade process. The Pitch Detection process finds similar waveforms with same pitch and overlaps , and adds up new wave form to modify the duration. Figure 9 shows how audio sound is prolonged in PSOLA method. Here waveform 2 is similar to waveform 1, and these 2 sequence is considered as having the same pitch. Concatenating waveform 2 with waveform 1 and overlapping the intersection, waveform 1 is then extended with the same pitch.

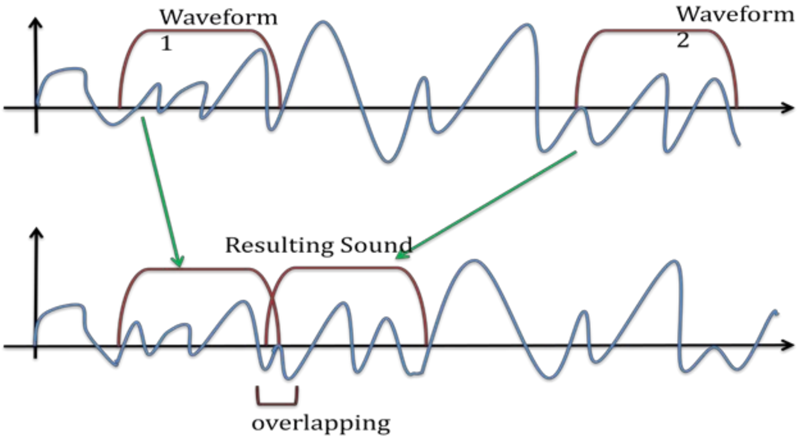

Figure 9. PSOLA Method Procedure to Prolong Audio

The traditional pitch detection algorithm used in PSOLA detects the maximum point of signal's auto correlation. The pitch detection process captures sections of waveforms from orignal audio clips and calculates each waveform's auto correlation. The two sections of waveform with the maximum correlation are considered as having the same pitch. Overlapping them with different overlapping length changes the duration of the original audio, and maintains the same pitch.

PSOLA's power consumption is extremely high in a real-time application such as slow motion video. Mediatek simplifies the pitch detection algorithm to achieve balance between CPU loading and audio quality. To reduce calculation, we only perform auto–correlation every 8 samples, and after finding the maximum correlation point in this coarse process, we do another round of auto correlation calculation for every 8 samples around the first maximum correlation point. In this way we can reduce CPU loading and still have a good enough quality.

Adding our time stretch technology along with Mediatek's slow motion video generates a slow motion video with audible sound, which is sync with video and has correct pitch.

Conclusion

With our powerful HW and optimized SW, Mediatek's Slow Motion Video technology provides a wonderful feature with excellent performance. Slow Motion Video makes the recording become more interesting and fun for the user. What's much more important, our business partners can create many more innovative features by using our solution.