Time Stamp Synchronization in Video Systems

Hsueh-szu Yang, Senior Software Engineer
Benjamin Kupferschmidt, Technical Manager

Teletronics Technology Corporation

Abstract

Synchronized video is crucial for data acquisition and telecommunication applications. For real-time applications, out-of-sync video may cause jitter, choppiness and latency. For data analysis, it is important to synchronize multiple video channels and data that are acquired from PCM, MIL-STD-1553 and other sources. Nowadays, video codecs can be easily obtained to play most types of video. However, a great deal of effort is still required to develop the synchronization methods that are used in a data acquisition system. This paper will describe several methods that TTC has adopted in our system to improve the synchronization of multiple data sources.

Keywords

Video Compression, Codec, Time Synchronization, Data Analysis, Real-Time Video, MPEG-2, MPEG-4, H.264, JPEG-2000

Introduction

MPEG-2 Transport Stream is a well-defined technology for transmitting or recording video data. It can be used with many different video compression standards such as MPEG-2, MPEG-4, and H.264. Multiple video and audio streams are allowed in a single transport stream. However, most applications use only one video stream and one audio stream. Several types of time stamps are also defined in the standard. Although it is optional to include time stamps in the stream, most implementations will include at least one or two time stamps in order to provide for synchronization between the audio and the video.

One of the most important data analysis tasks that must be performed on a data acquisition system is the correlation of video frames with events, sensor data, avionics bus data or other video frames that occur at the same time. Unfortunately, the time stamps provided by MPEG-2 Transport Streams are insufficient for synchronizing most data acquisition systems unless additional information is provided. In some implementations, a sync marker is simultaneously inserted into all data and video channels. TTC's method works a different way; it takes advantage of accurate

synchronized time information, which can be obtained from IEEE-1588 or IRIG time. Since a time synchronization mechanism already exists in the data acquisition system, we do not need to insert a sync event into the data.

JPEG-2000 video stream is another approach that TTC implemented. IRIG time stamps are inserted into a private header. It is a simpler and direct way to make each frame have the time stamp which share a common base time with the other channels. Therefore, synchronization will not be an issue.

The next section of this paper will explain how time stamps can be used to synchronize video with other data channels that occur on a time line. After that, we will discuss several issues related to time synchronization and the results from testing that we performed with our time synchronization method. Finally, we will describe three complete time synchronized video systems.

Video Synchronization Challenges

Speed Control

In some simple implementations, there might be no need to control the speed that video plays. These simple implementations rely on the video source to send video frames at a constant rate. The video frames are displayed as soon as they arrive. An example of this is a traditional television broadcast. In a TV broadcast uncompressed video data is transmitted at a constant bit rate. This has very little delay but requires a large amount of bandwidth.

The story is different for compressed video. Speed control is usually required on the playback side. One reason for this is that the size of each frame is very different for MPEG-2, MPEG-4, and H.264. The three most commonly used frame types are: I-frame, P-frame and B-frame. Usually they are interleaved into the video stream. Among the three, the I-frames are the largest. They require more bandwidth and a longer time to transmit. The B-frames are the smallest. For constant bit rate transmission, the video decoder will spend more time waiting for I-frames then it spends waiting for P-frames or B-frames. As the result, the time when each frame is received is not the correct time for displaying the frame. Even if we use 100 Megabit Ethernet or higher bandwidth, the nature of Ethernet and TCP/IP is non-deterministic with regard to transmission time. This means that the arrival time will be highly variable depending on the network status.

Moreover, if B-frames are used in the compression algorithm, then the transmission frame order is different than the order that the frames were captured. For example, a B1 frame is captured earlier than a P2 frame, but it will be transmitted after the P2 frame. This reordering of the video data means that speed control is required in order to show each picture at the right time with the right duration.

To implement speed control, we will require time information in the video stream. The implementation of speed control will be discussed in the next section.

Jitter and Choppy Video

Jitter and choppy video is a symptom that a video picture is not being displayed at the right time and for the right duration. It can also indicate that a significant number of frames are being dropped and that the movement in the video will not be smooth. These symptoms also imply that the frame rate is not constant. There are many possible causes of this problem. For example, when the system is overloaded, the CPU or DSP cannot encode or decode a frame in time. This can cause transmission delay or it can cause the pictures to show up late on the display. This will result in slow playback. Later when the system is trying to catch up, it will play faster than normal. Another possible cause of this problem is internal buffer underrun or overflow. Temporary lack of sufficient bandwidth is yet another reason.

There is always latency between the video encoder and decoder. The latency is a sum of the encoding delay, the decoding delay, the transmission delay, and the propagation delay. The result of the delays will result in jitter and choppy video. We can statistically get deviation values, such the average deviation and the maximum deviation, from the time stamp information provided in the video stream.

The deviation value can also be used to adjust the presentation time of each frame. This will result in smoother video, assuming that there are no dropped frames. If choppiness is caused by dropped frames, it cannot be fixed. Although delaying the presentation of the video will also smooth the video, a long delay is generally undesirable for real-time video playback. The software must balance a longer delay versus less jitter and choppiness. In any case, the system implementation needs to minimize the deviation. Otherwise the deviation might be too large which would make it impossible to compensate for. This would lead to unavoidable jitter and choppy video playback.

Utilizing Time Information

Timestamp and Reference Clock

The concept of a speed control is simple: if there is a time stamp for each video frame and there is a reference clock then the video player just needs to read the time stamps and wait until the right time to put each frame on the display. However, there are many variables involved in implementing this feature.

If there is no timestamp in the MPEG video stream but we are provided with a frame rate then a time stamp can be calculated for each video frame based on the frame rate. For the reference clock, the player can use a local timer. This example is simple, but there are several issues. The playback speed of each video or audio channel will be correct.

However, there is no way to synchronize between the video, audio and other channels of data. The time stamp that is derived from the frame rate is essentially a private time. It does not share a common base time with the other channels. This means that it is vulnerable to lost data or dropped frames. The frequency drift in the crystal oscillators will result in that the speed of encoder side being slightly different than the decoder side. This will cause a problem for real-time video applications after they have been running for a while.

In the MPEG video system, there are several types of timestamp information that can be used to overcome these problems:

    1. PTS in the PES syntax
    2. DTS in the PES syntax
    3. PCR in the transport stream syntax
    4. OPCR in the transport stream syntax
    5. SCR in the program stream syntax
    6. ESCR in the PES syntax

PCR, SCR and ESCR are reference time stamps. They provide information for adjusting the reference clock. In some circumstances, they can be used as a reference clock.

The PTS (Presentation Time Stamp) is the exact time that a frame needs to display. Most MPEG-2, MPEG-4 or H.264 streams have PTS. Video, audio and data in the same program stream must use the same base time. Therefore, the synchronization between channels in the same program stream can be achieved by using this time stamp. DTS stands for Decoding Time Stamp, which indicates the time that a frame should be decoded.

PCR (Program Clock Reference) represents a system time clock on the encoder side. Depending on the codec implementation, this piece of information can be used for determining the initial value for a reference clock or to estimate the jitter. If it is used to estimate the jitter then it can be used to determine the buffer size and the delay in the decoder side. For a real-time application, those parameters will decide the smoothness of playback and the latency.

GPS Time

The standard only requires using the same base time when the video, audio and data streams are in the same program stream. A GPS time stamp can be used for synchronization of data in different program streams. It can also be used to synchronize data that is encoded or recorded by different units.

In most of data acquisition systems, there is a time source. The time source is usually synchronized with GPS time. Different units in the data acquisition system can use IEEE-1588 or IRIG time input to accurately synchronize with the time source. This provides

the video units a way to include extra time information. The extra time information can be used later to convert the time stamp from PTS format to GPS time.

This implementation is very similar to Chapter 10 time packets. The time packet provides information so that the time stamp for each video picture will have the same base time. This will make it easy to synchronize the video pictures, MIL-STD-1553 bus data, events and sensor data from a PCM stream. This can be very accurate for correlating video images with data.

The frequency drift in the crystal oscillator can also be resolved by using GPS time on both the encoder and the decoder side. Instead of using the counter from the crystal oscillator, the GPS time stamp of each frame can be used for speed control. The difference between the frame rate on the decoder side and the encoder side will be very small. This will eliminate the delay that occurs when the decoder runs slower than the encoder. It also eliminates choppy playback and buffer underruns that could occur when the encoder runs too fast.
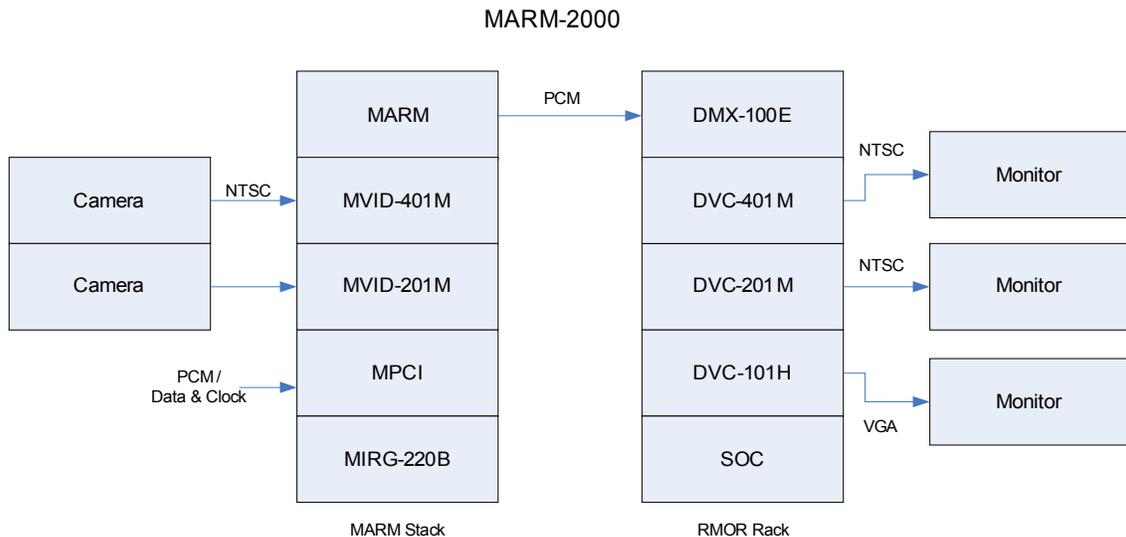
Examples

Figure 1: Video and Data Telemetry Using MARM-2000 / RMOR-2000

In this example, video and PCM data are acquired by a MARM-2000 unit. A composite PCM stream will be created and sent to the RMOR-2000 Rack Mounted Reproducer. The RMOR-2000 decodes the video and regenerates the PCM data in real-time.

The MIRG-220B is the time module, which is able to accept IRIG time. The MVID-401M and MVID-201M are video encoder modules that accept standard video signals and compress the data into MPEG-4 and MPEG-2 video streams respectively. Time information from the MIRG-220B will be included in the video streams. The MPCI-102

modules can accept any standard PCM stream. This PCM stream could include time information inside. All of the PCM input and video streams are multiplexed together into a composite PCM stream and transmitted.

On the RMOR-2000 side (the ground demultiplexer side), the DMX-100E card accepts the composite PCM stream and dispatches each embedded data channel to the corresponding decoder board in the rack mounted unit. The DVC-401M, DVC-201M and DVC-101 cards are able to extract the video streams from the composite PCM, decode it and output it to a monitor in real-time.

Time information in the video stream will be utilized by the DVC boards. This helps to ensure that there is accurate synchronization between the video and audio data. It also helps to select a buffer size for each DVC board so that the video output is smooth.

JPEG-2000 Video

MCVC-501J and MVID-501J are JPEG-2000 video encoders. They insert a private header, in which there is an IRIG time stamp. Unlike MPEG video, there is not extra step needed to convert the time stamp. The frame order is same during transmission and when frames capturing. DVC-101J can be used as the decoder. JPEG-2000 has only one type of frame. The deviation of delay is small. It does not require information to estimate the deviation of delay, such as PCR used in MPEG-2 transport stream.
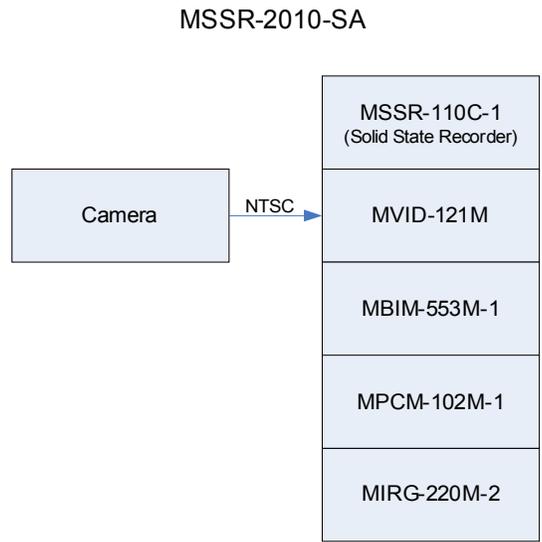
Video and Data Recording

MSSR-2010-SA

Figure 2: Video and Data Recording In Chapter 10

The MSSR-2010-SA unit has a solid state recorder that can record multiple channels in Chapter 10 format. The MIRG-220M-2 module provides time information for all of the data acquisition modules in the unit. The MVID-121M module accepts a standard video

input and encodes the video data. The MBIM-553M-1 is a MIL-STD-1553 bus monitor and the MPCM-102M-1 accepts two PCM input streams.

The MVID-121M is able to generate an MPEG-2 stream with extra time information. This extra information can be used to convert PTS time to match the system time that is obtained from the MIRG-220M-2. The MSSR-110C-1 recorder will insert time packets based on the time stamps from the MIRG-220M-2 module. The time stamp for each video image and the time stamp in the Chapter 10 header use the same base time. In data acquisition systems, each MIL-STD-1553 bus can have a time stamp. The PCM streams also have their own time stamps. Since all of the data channels receive their time stamps from the same source, it is possible to correlate 1553 events, PCM data and video images with the main Chapter 10 time stamp.
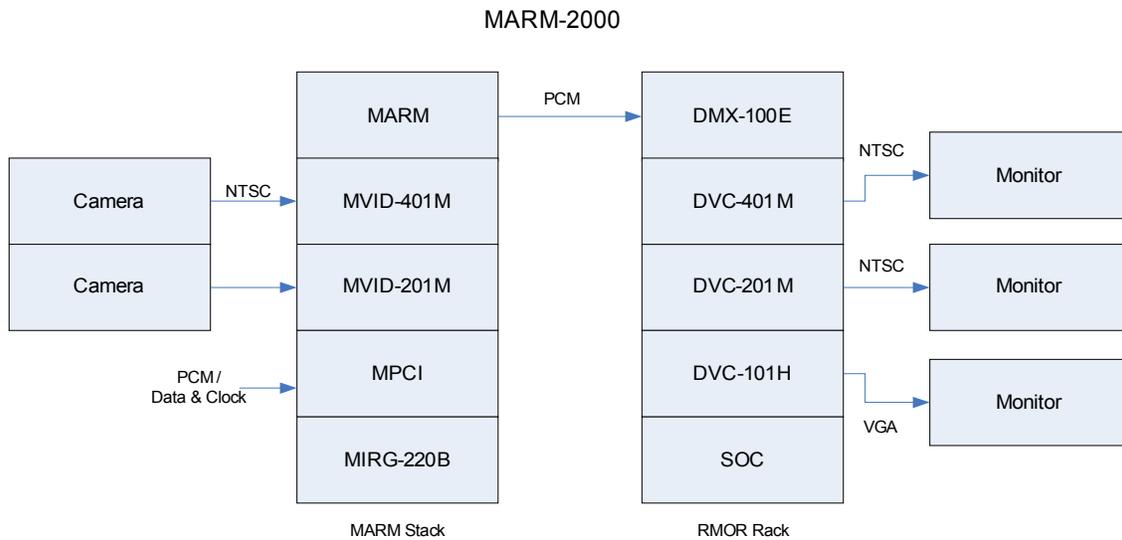
MARM-2000

Figure 3: JPEG-2000 Using MARM-2000 / RMOR-2000

Conclusion

Time stamps are essential information for synchronization between video pictures, events and sensor data. PTS (Presentation Time Stamp) and PCR (Program Clock Reference) are usually used for synchronized between video and audio in MPEG-2, MPEG-4 and H.264. Extra time information is included into the video stream in TTC's solutions. This extra time information describes the relationship to IRIG time or GPS time. With the extra time information, all data channels and video channels can share the same common base time. This means that we can correlate all of the time stamps together.

JPEG-2000 provides a simpler but robust way to synchronize with other data channels. Each picture header has a private header that contains an IRIG time. No extra step is required to convert the time stamp in order to synchronize with other channels. There is

no inter-frame compression and frame order is not re-ordered. All these can significantly simplifier the data analysis process if compression efficiency is not a major concern.

The PCR and PTS are also used to control the delay and they result in smoother video playback for real-time video applications. In the TTC solution, the PCR and PTS are set to compensate for the deviation of the delay value. This helps to create smoother video. The playback delay is controlled as a constant value. This is different then in other implementations where the delay is constantly accumulating and growing. A constant delay is more suitable for real-time monitoring.