Module 10

MULTIMEDIA SYNCHRONIZATION

Lesson 35

Time stamping and pack architecture.

Instructional Objectives

1. Specify the requirements of media playback.
2. Define packets and packs for media stream
3. Present the block diagram of Digital Storage Medium (DSM) and System Target Decoder (STD)
4. Specify the requirements of systems clock reference (SCR)
5. Show the pack architecture
6. Define the elements of the pack header.

35.0 Introduction

In lesson-34, we have studied the reference model for multimedia synchronization and the synchronization requirements in a distributed environment. The issue of synchronization has been addressed in MPEG-2 standard, where intra-media continuity and inter-media synchronization have been handles at different layers of the multimedia stream. In this lesson, we are going to study the architecture of multimedia streams and the time-stamping requirements to ensure synchronization between the digital storage medium (DSM) , i.e. the source and the systems target decoder (STD), i.e. the sink.

35.1 Requirements of media playback

There are two distinguishing requirements of media playback :

- Intra-media continuity
- Inter-media synchronization

In order to understand MPEG synchronization, it is necessary first to familiarize with the encoder's system architecture. At the time of encoding in MPEG, in the first stage, video frames and audio samples are separately encoded, and an independent stream of bytes is output for each media channel. Each media stream is then packetized independently. Packets from the different streams are interspersed to form a single multiplexed stream of packets, and the multiplexed stream is then organized into packs, with each pack comprising of an integral number of packets.

35.2 Packs and packets

In a MPEG stream, whereas continuity is handled at the pack layer, synchronization is handled at the packet layer. Fig.35.1 shows the generation of packs at the encoder. The decoder system referred to an System Target Decoder (STD), as shown in fig.35.2.

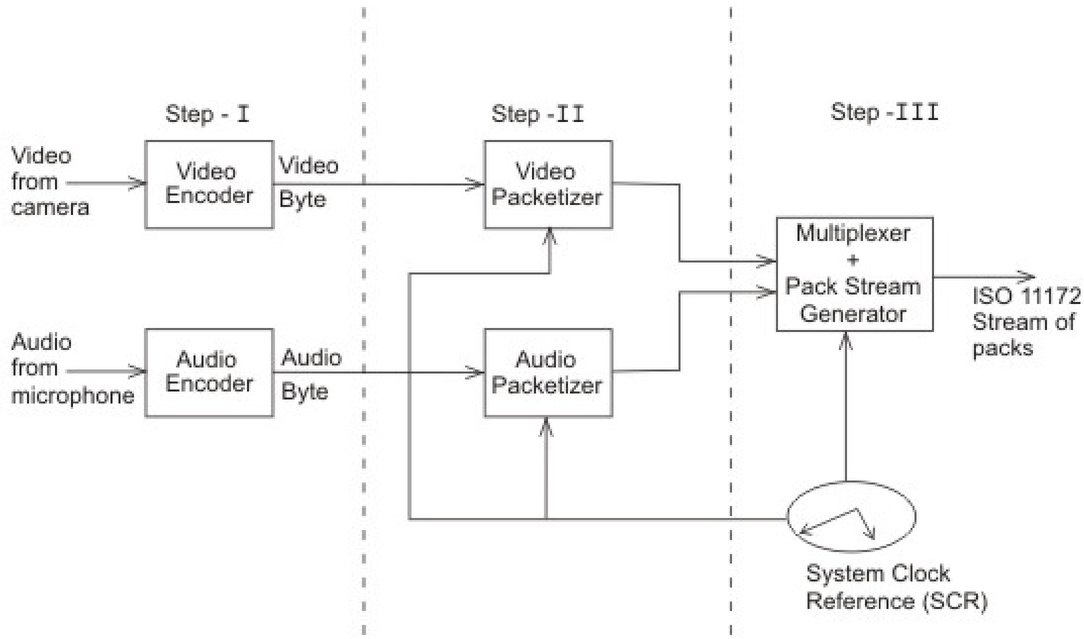

FIGURE 35.1 Generation of packs at the encoder

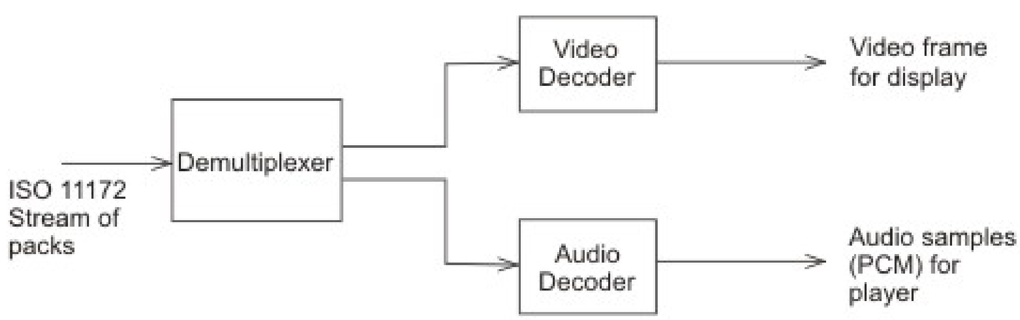

Figure 35.2 System target decoder

35.3 Requirements of System Clock Reference (SCR)

In MPEG stream, packs constitute a common system layer appeared around a media specific compression layer, and house the following functions at the encoder :

- Interleaving of multiple compressed streams into a single stream
- Time stamping : insertion of the SCR

SCR's are sampled values (in Hz) of the encoder's clock which is constrained to meet the following drift bounds on its frequency :

- Nominal frequency = 90,000 Hz
- Fractional frequency drift $\rho \leq 0.00005$
- Frequency drift $\upsilon \leq 4.5$ Hz
- Rate of change of drift $\delta\upsilon/\delta t \leq 0.00025$ Hz/s.

Taking drift at the rate of 0.00025 Hz/s, the frequency can go beyond the permitted 4.5 Hz drift bound over a period of about 5 hours. Hence, continuous operation of the encoder for longer than 5 hours may need resynchronization. SCR is inserted into each pack header. The SCR value in a pack is set equal to reading of the encoder's clock at the instant the last byte of the SCR is constructed at the encoder. Successive SCR's cannot differ by more than 0.7 seconds to assure accurate clock updating at the STD.
Hence, the time interval between successive packs cannot exceed 0.7 seconds.

A pack header also contains the rate of the multiplexed stream as manifested in the pack and is termed the *mux-rate*. The *mux-rate* is not the bit or the byte rate but is a scaled value.

Suppose, the number of bytes following the SCR in a pack p till end = l_p bytes and t_l is the time at which the last byte of the pack is constructed at the encoder. Then, the *mux-rate* is computed as

$$mux-rate = \frac{l_p}{50\left(t_l - SCR\right)}$$

The *mux-rate* may vary from pack to pack.

The first pack header usually contains the following additional system headers

- ***Rate bound*** **:** Max. *mux-rate* coded in any pack.
- ***Video bound:*** Integer, indicating the maximum number of video streams (upto 16).
- ***Audio bound:*** Integer, indicating the maximum number of audio streams (upto 32)
- ***Video, Audio lock flag :*** indicates if there is a specific constant harmonic relationship between the media sampling rates and the SCR frequency.
- ***Fixed flag*****:** defined in the following section.

35.4 Pack architecture and pack headers

Fig.35.3 shows the composition of a pack from the packets.

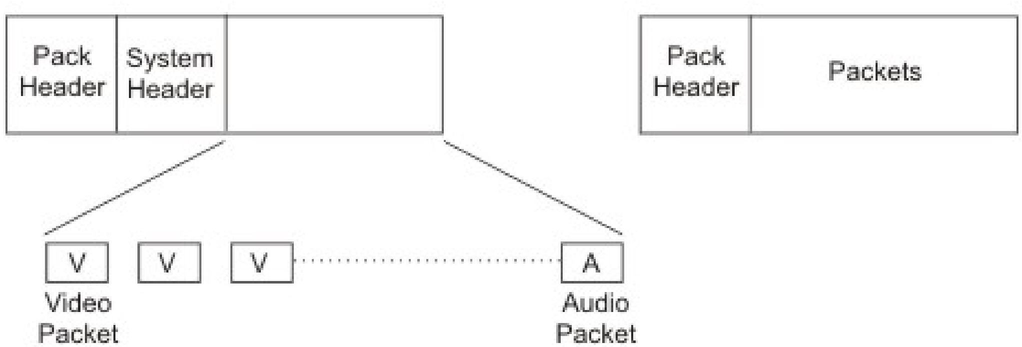

Figure 35.3 Composition of a pack

The *fixed flag* indicates whether the stream is of fixed bitrate, in which case, SCRs of each pack in the multiplexed stream bear the following linear relationship with the size offset of the pack within the stream.

$$SCR(i) = \lfloor (a * i + b) \rfloor \text{ modulo } 2^{33}$$

where *a* and *b* are real-valued constants, *i* is the index of the final byte of the SCR field counted from the start of the multiplexed stream. If the data rate averaged between the successive SCRs is equal to the *max-rate* field, the *fixed flag* is set.

If the *mux-rate* field indicates higher than the average rate, then the stream is delivered in bursts. In order to transform a bursty stream into a constant bit rate stream, MPEG permits an encoder to insert a padding stream. Such a padding stream would consist of a sequence of packets interleaved with audio and video packets within each pack to achieve a constant bit rate over the entire pack. At the STD upon demultiplexing, the padding stream packets are simply discarded.

35.5 Conclusion

In this lesson, we have studied the requirements of inter-media synchronization using time-stamping for media playback. The media stream is broken into packets, containing only one type of media objects and packs, which combine several packets. We have presented the pack architecture in this lesson. In the next lesson, we are going to present the architecture of packets, the interleaving of packets of different media streams and the intra-media continuity aspects in playback.