Module 10

MULTIMEDIA SYNCHRONIZATION

Lesson 33
Basic definitions and requirements

Instructional objectives

At the end of this lesson, the students should be able to:

1. Define synchronization between media objects.
2. Distinguish between time independent and time dependent media objects.
3. Define continuous media objects.
4. Define intra-object synchronization and inter-object synchronization
5. Define Logical Data Units (LDU) with examples.

33.0 Introduction

In the previous lessons, we have discussed at length about some of the major components of multimedia communication, like images, video and audio. A multimedia presentation is always incomplete unless the individual media streams composing the presentation are presented in a meaningful way. This is addressed by multimedia synchronization, the topic of the current module.

The present lesson first outlines some of the fundamental concepts and definitions related to multimedia synchronization. This is followed by a four layer synchronization reference model, which is needed to understand the various requirements for the run time mechanisms, identify interfaces between run time mechanisms and comprising the system solutions

33.1 Synchronization between media objects

By now we clearly understand what is a media object, but we are required to understand what relationships should exist between the media objects in a multimedia presentation. The relationships may be temporal, spatial or content based.

The word *synchronization* refers to time, but in a broader sense, the definition of synchronization could be extended to content or the spatial relationships.

In a strict sense, multimedia synchronization would mean temporal relationships between the media objects. A common example of temporal relationship is movie or television, where both audio and video objects are involved. There it is essential to maintain the same temporal relationship, as it would exist for recording, during the replay too.

A typical example of content relationship is a spreadsheet and a graphics that correspond to the spreadsheet data. In this case, the same data are represented in two different ways.

The spatial relationships refer to the space used for presenting a media object to an output device at a certain point of time. It therefore defines where in space (usually the two dimensional display screen) should the media objects be placed.

33.2 Time-independent and time-dependent media objects

Before we study the mechanisms of temporal synchronization between the media objects, we should understand the difference between time independent and time-dependent media objects.

We may note a fundamental difference between how an image or a text is transmitted and how a video or an audio is transmitted. In the former type, we may transmit and update the image or text from left to right, or right to left, or top or bottom (the usual way!) or the bottom to top. In the reconstructed presentation, the semantics of the content do not change whichever way we do, provided the same order is maintained at the transmitter and the receiver. There may be pauses while updating, but still nothing matters to the semantics. However, for a video or an audio, the order is important. We cannot run a video or an audio presentation in the reverse order, since the semantics will be lost. If a pause is introduced, annoying effects, loss of intelligibility (for audio) would result. From these considerations, we can state the difference between time-dependent and time-independent media objects.

A *time-dependent media object* always has a media stream associated with it. The media stream can be divided into units and temporal relationship exists between the units of the media stream.

A *time-independent media object* is a traditional media object like image, graphics or text. The semantic of the media content does not depend upon a presentation in time domain.

A time-dependent media object is said to be a *continuous media object* if the presentation durations of the unit of the media stream are all equal. For example, video is a continuous media object, since all units (frames) are of equal durations.

33.3 Intra and inter-object synchronization

Synchronization requirements exist within one time-dependent media object and also in between several media objects.

Intra object synchronization refers to the temporal relationship between various presentation units of a time-dependent media object. For example, if a video presentation has a refresh rate 30 frames/ second, then each frame should be

displayed for 33 ms. Otherwise, the synchronization will be lost. Lack of intra object synchronization results in pauses or jerks.

Inter-object synchronization refers to the temporal relationship between several media objects, which may be time-independent or time dependent fig.33.1 shows an example of inter media synchronization. In this example, we first have a video, which has a synchronized audio with it. This is followed by several images, which is again followed by an animation, synchronized with audio commentary.

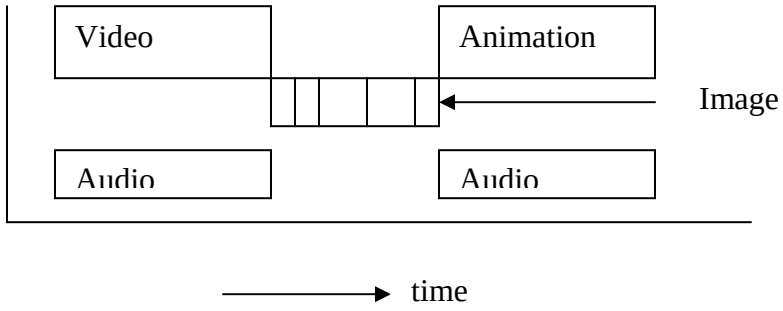

Fig. 33.1. Example of inter-media synchronization.

If in fig 33.1, we add to temporal offset to the audio with respect to the video or the animation, a lack of inter object synchronization results.

33.4 Logical Data Unit (LDU)

We now make a clear distinction between the presentation units and logical data units of a time dependent media object. For example, in a video sequence, each frame always represents a presentation unit. The logical data unit (LDU) on the other hand refers to an information unit, which may not be the same as that of the presentation unit. How we define the information unit depends upon the level of granularity. Each frame consists of macroblocks of size 16 x 16 pixels. Each macroblock may compose an LDU. If we increase the level of granularity further, each pixel may also form an LDU.

Each LDU may not contain the same amount of information. For example, if we consider the compressed frames of a video sequence as LDUs, each compressed frames may require different number of bits to encode.

LDUs may also classified into closed and open LDUs. *Closed LDUs* are the ones having predictable duration, like the stored continuous media, e.g. audio and video. *Open LDUs* typically represent input from a live source.

Some typically examples of LDU based syncronization are shown below:

- Lip synchronization, which requires tight coupling between audio and video, as illustrated in fig 33.2

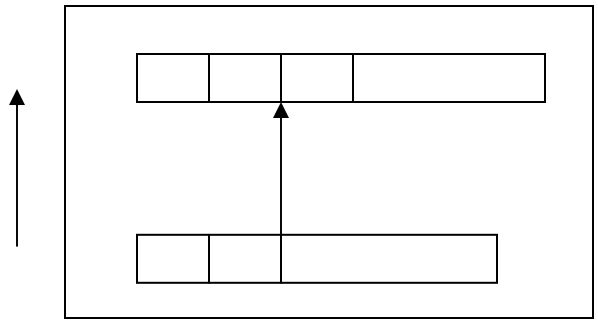

Fig 33.2 LDU view of lip synchronization

The synchronization is specified in terms of maximal skew between the video and the audio streams.

- *Slide show with audio commentary*, which demands that the change of slides should be temporally related to the audio commentary.

- *Synchronized multimedia presentation through interactions*: This may be done in a variety of ways. For example, a synchronized audio/ video presentation could be followed by an interactive session, where the user may exercise own choice to view a set of slides or another synchronized presentation, followed by interactions etc.

Source: http://nptel.ac.in/courses/Webcourse-contents/IIT%20Kharagpur/Multimedia%20Processing/pdf/ssg_m10l33.pdf