CS 414 – Multimedia Systems Design

Lecture 33 – Synchronization (Part 1)

Klara Nahrstedt

Spring 2009

Administrative

- MP4 posted
  - April 30 (preview for finalists) 5-7pm
  - May 1 grading for non-finalists 3:30-5pm and competition for finalists 5-7pm
- Peer Evaluation
  - Deadline May 6, 5pm – very important 5% of your grade
  - Website will have all the instructions about the peer-evaluation (see also the first lecture of the semester that discussed peer evaluation)

Outline

- Synchronization Concept
- Synchronization Classification
- Logical Data Units
- Live vs Synthetic Synchronization
- Synchronization Requirements

Notion of Synchronization

- Sync in correspondence to
  - Content relation
  - Spatial relation
  - Temporal relation
- Content Relation
  - Define dependency of media objects for some data
  - Example: dependency between spreadsheet and graphics that represent data listed in spreadsheet

Spatial Relation

- Layout relation
  - Defines space used for presentation of media object on output device at certain point of multimedia presentation
  - Example: desktop publishing
- Layout frames
  - Placed on output device and content assigned to frame
  - Positioning of layout frames:
    - Fixed to position of document
    - Fixed to position on page
    - Relative to position of other frame
  - Example: in window-based system, layout frames correspond to windows and video can be positioned in window

Temporal Relation (Our focus!!!)

- Defines temporal dependencies between media objects
- Example: lip synchronization
- **Time-dependent object**
  - Media stream since there exist temporal relations between consecutive units of the stream
- **Time-independent object**
  - Traditional medium such as text or images
- **Temporal synchronization**
  - Relation between time-dependent and time-independent objects
  - Example: audio/video sync with slide show

Temporal Relations

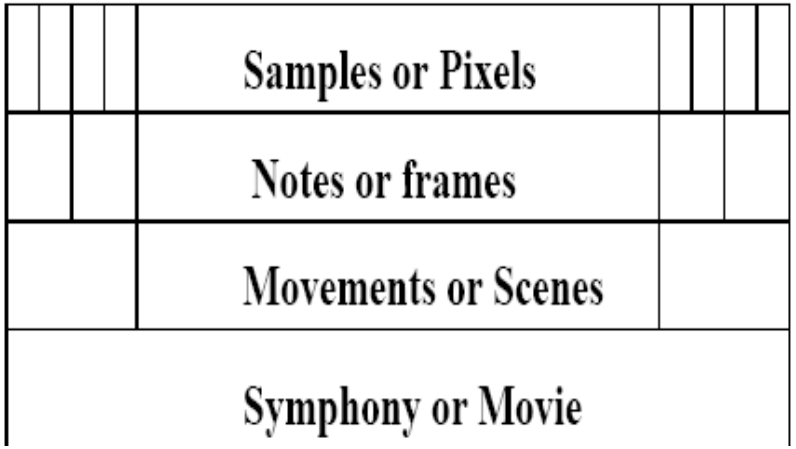
- Synchronization considered at several levels of Multimedia Systems
- **Level 1: OS and lower level communication layers**
  - CPU scheduling, semaphores during IPC, traffic shaping network scheduling
  - Objective: avoid jitter at presentation time of one stream
- **Level 2: Middleware/Session layer (Run-time)**
  - Synchronization of multimedia streams (schedulers)
  - Objective: bounded skews between various streams
- **Level 3: Application layer (Run-time)**
  - Support for synchronization between time-dependent and time-independent media together with handling of user interaction
  - Objective: bounded skews between time-dependent and time-independent media

Synchronization Specification

- Implicit
  - Temporal relation specified implicitly during capturing of media objects
  - Goal: use this temporal relation to present media in the same way as they were originally captured
  - Example: Audio and Video recording and playback
- Explicit
  - Temporal relation specified explicitly to define dependency in case media objects were created independently
  - Example: creation of slide show
    - Presentation designer
      - selects slides,
      - creates audio objects,
      - defines units of audio presentation stream,
      - defines units of audio presentation stream where slides have to be presented

Logical Data Units and their Classification

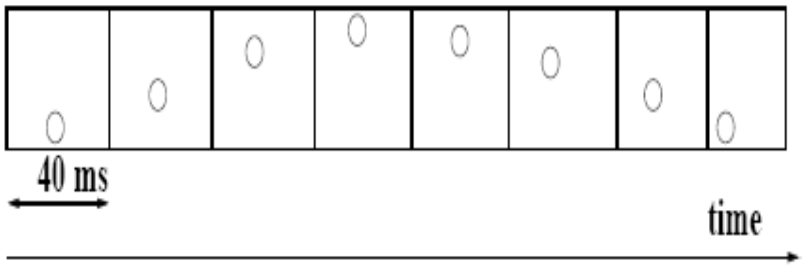
- Time-dependent presentation units are called logical data units (LDU)s.
- LDU classification
  - Open
  - Closed
- LDUs important
  - In specification of synchronization



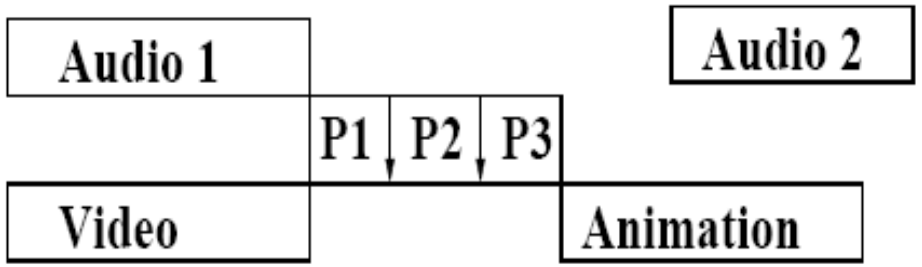
	LDU spec during capturing	LDU spec defined by user
Fixed LDU	Audio/Video	Animation/Timer
Variable LDU	Recorded Interactions	User Interactions

Synchronization Classification

- Intra-object Synchronization
  - Time relation between various presentation units of one time-dependent media stream



- Inter-object Synchronization
  - Time relation between media objects belonging to two dime dependent media streams



Synchronization Classification

- Live Synchronization
  - Goal: exactly reproduce at presentation temporal relations as they existed during capturing process
  - Requirement: must capture temporal relation information during media capturing
  - Example: video conference, phone service
  - Example: recording and retrieval services – presentations with delay

Synchronization Classification

- Synthetic Synchronization
  - Goal: arrange stored data objects to provide new combined multimedia objects via artificial temporal relations
  - Requirements: support flexible synchronization relations between media
  - Example: authoring, tutoring systems
- Two phases:
  - Specification phase – define temporal relations
  - Presentation phase – present data in sync mode
- Example:
  - 4 audio messages recorded related to parts of engine in animation. Animation sequence shows a slow 360 degree rotation of engine

Synchronization Requirements during media presentations

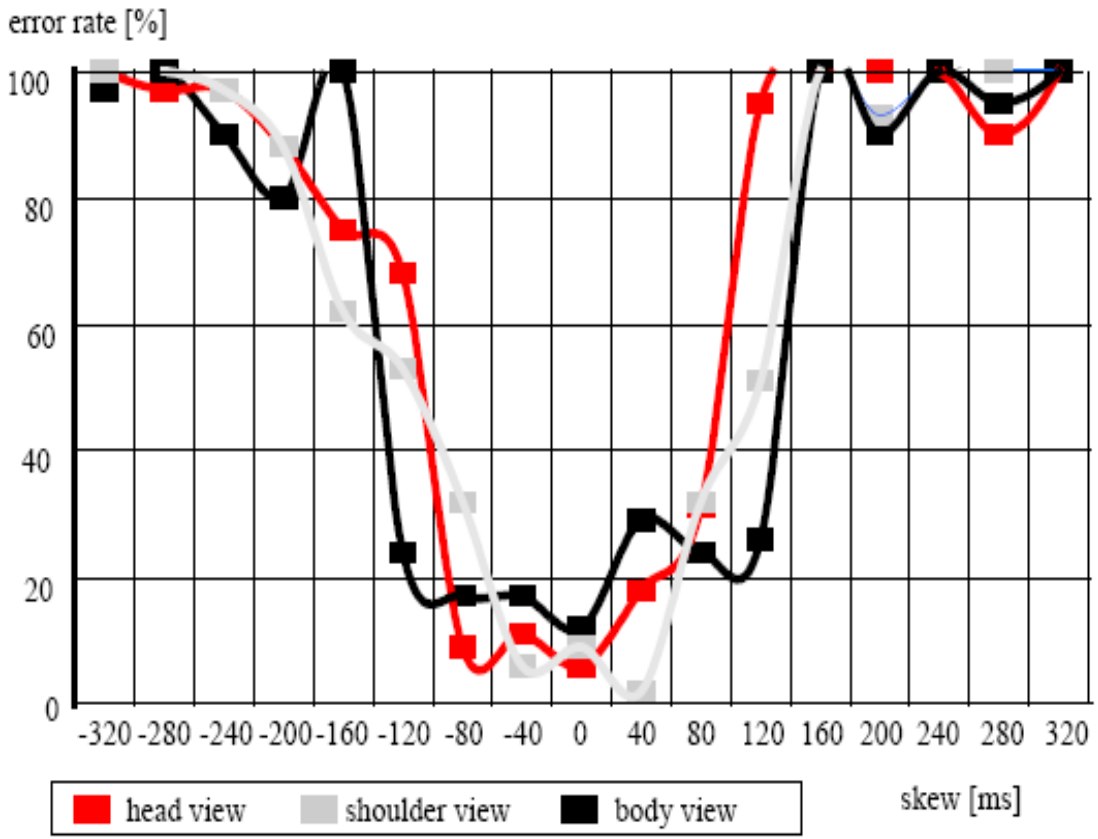
- For intra-object synchronization
  - Need accuracy concerning jitter and EED delays in presentation of LDUs
- For inter-object synchronization
  - Need accuracy in parallel presentation of media objects
- Implication of blocking:
  - O.K. for time-independent media
  - Problem for time-dependent media – **gap problem**

Gap Problem in Synchronization

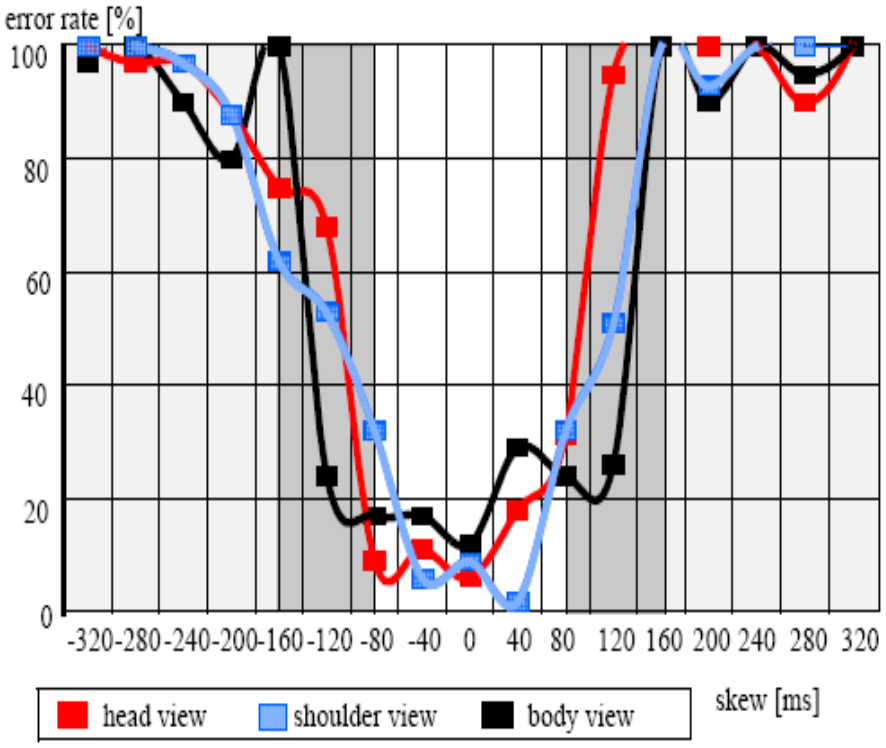
- What does blocking of stream mean for output device?
  - Should we repeat previous music, speech, picture?
  - How long should such gap exist?
- **Solution 1: restricted blocking method**
  - Switch output device to last picture as still picture
  - Switch output device to alternative presentation if gap between late video and audio exceeds predefined threshold
- **Solution 2: resample stream**
  - Speed up or slow down streams
  - Off-line re-sampling – used after capturing of media streams with independent streams
    - Example: concert which is captured with two independent audio/video devices
  - Online re-sampling – used during presentation in case gap between media streams occurs

Lip Synchronization

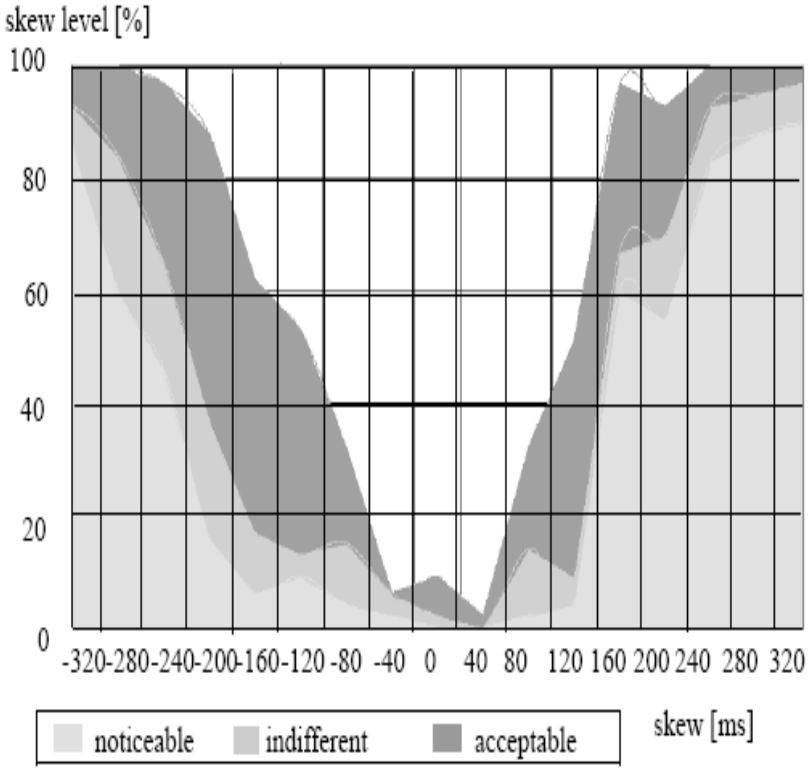
- Temporal relation between audio and video
- **Synchronization skew**
  - Time difference between related audio and video LDUs
- **Streams in sync** iff skew = 0 or skew ≤bound
- **Negative skew**: video before audio
- **Positive skew**: Audio before video



Lip Synchronization



Perception of Synchronization Errors

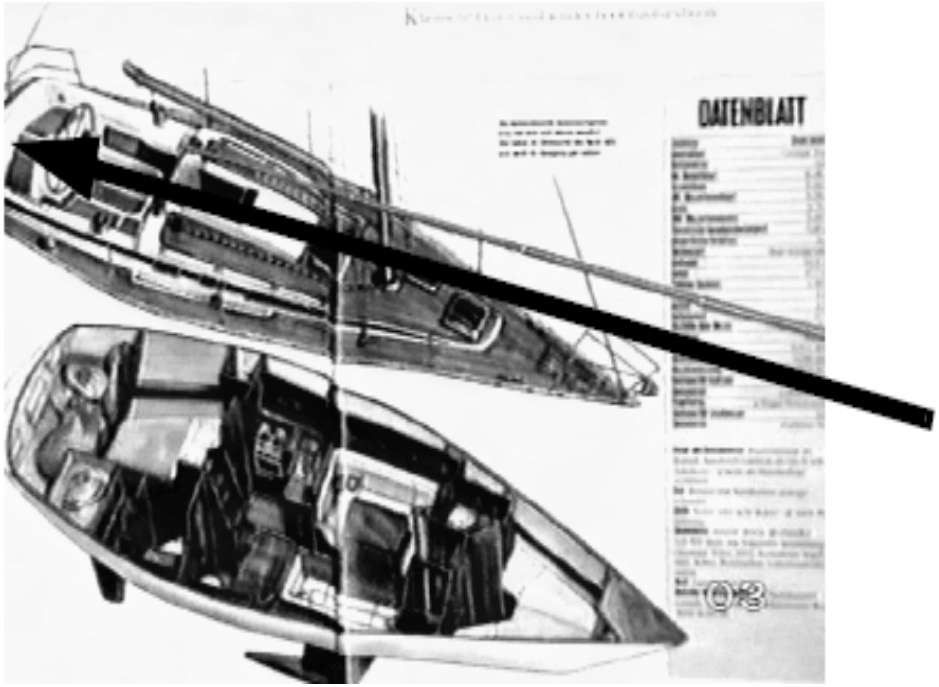


Skew Level found to be annoying

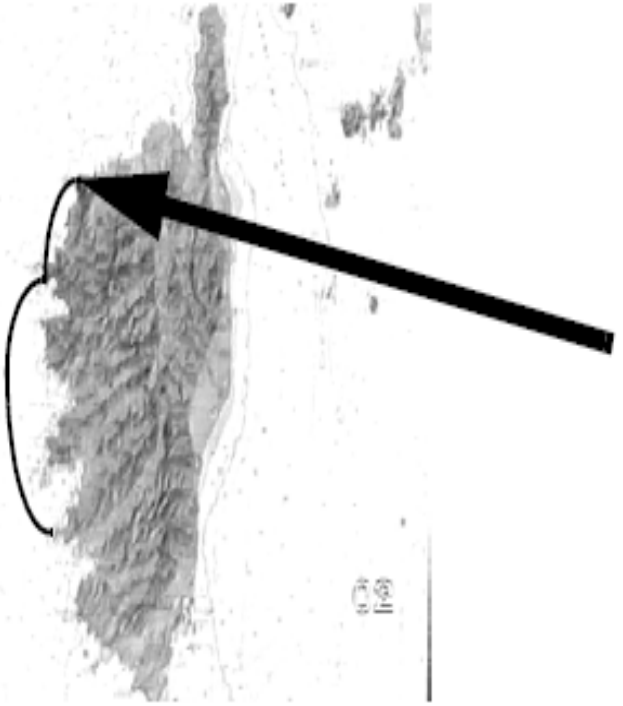
Lip Synchronization Requirements

- In sync:
  - -80ms ≤ skew ≤ 80ms
- Out of sync:
  - Skew < -160ms
  - Skew > 160ms
- Transient:
  - -160ms ≤ skew < -80ms
  - 80ms < skew ≤ 160ms

Pointer Synchronization

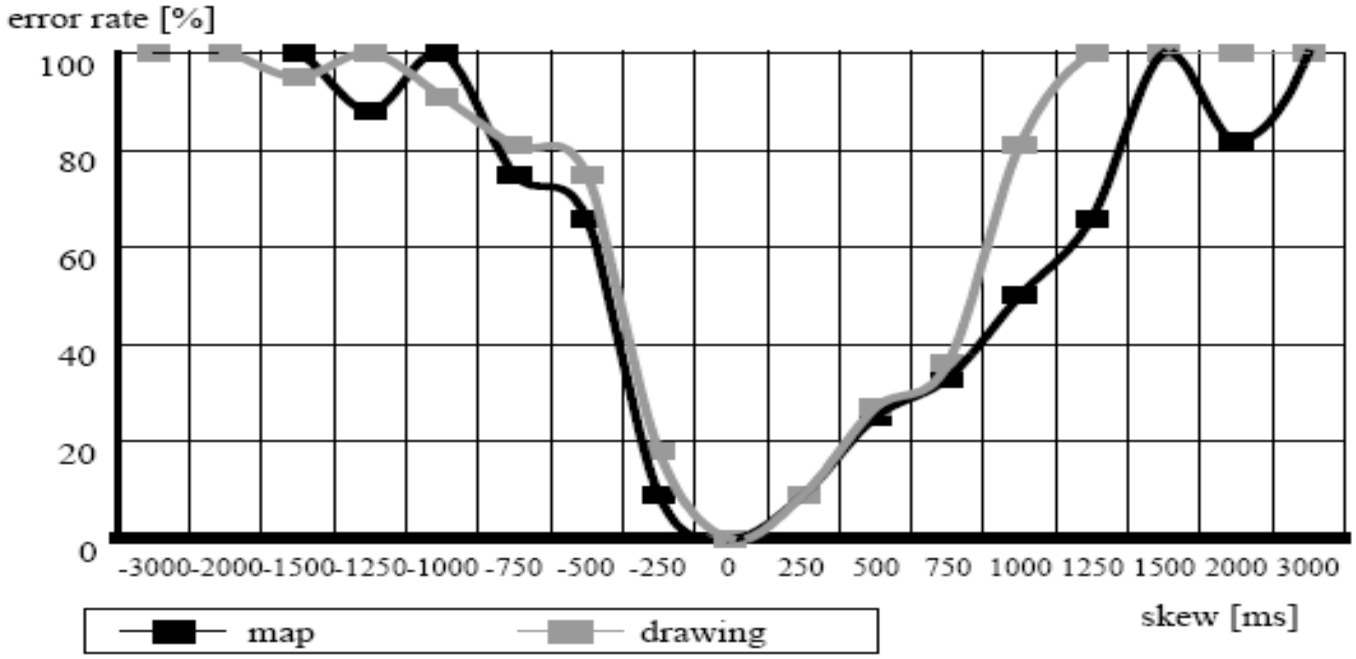


Pointer Sync based on technical drawing



Pointer Sync based on map

Pointer Synchronization



Negative skew: pointer before audio
Positive skew: pointer after audio

Pointer Synchronization Requirements

- In sync:
  - -500ms ≤ skew ≤ 750ms
- Out of sync:
  - Skew < -1000ms
  - Skew > 1250ms
- Transient sync situation:
  - -1000ms ≤ skew < -500ms
  - 750ms < skew ≤ 1250ms

Other Sync Requirements

- Jitter delay of digital audio
  - Max. allowable jitter:
    - 5-10 ns (perception experiments)
    - 2 ms (other experiments)
- Combination of audio and animation
  - Not stringent as lip sync
  - Max allowable skew: +/- 80ms
- Stereo audio
  - Tightly coupled
  - Max allowable skew: 20 ms
  - Due to listening errors, suggestion even +/- 11ms
- Loosely coupled audio channels (speaker and background music)
  - Max allowable skew: 500ms

Conclusion

- Carefully analyze what kind of synchronization is needed in your multimedia system and application
- Determine at which level you need synchronization
- Determine what the synchronization requirements should be based on prior experiments