IDIAP COMMUNICATION

From Meeting Recordings to Web Distribution: Description of the Process

Maël Guillemot * Bastien Crettol *
Pierre Wellner *

IDIAP–Com 05-05

January 4, 2006

* IDIAP, Martigny, Switzerland

IDIAP Research Institute www.idiap.ch
Rue du Simplon 4 P.O. Box 592 1920 Martigny – Switzerland
Tel: +41 27 721 77 11 Fax: +41 27 721 77 12 Email: info@idiap.ch

1 Outline

This document aims at reporting in detail the steps needed from capturing a meeting to its distribution on the media file server. This can be structured as three main phases. Phase 1 - the recording equipment setting before a meeting, including instructions to participants. Phase 2 - the set of media (audio, video, slide captures, pen and whiteboard strokes) processing to run after a meeting has been recorded. Phase 3 consists of publishing the secured web distribution of the data. The different steps within each phase are described in order, with some notes on what can be performed in parallel. Also we list current issues and stating what could be improved to maximize the automation of all this process.

2 Introduction

Research on meeting recordings is conducted world-wide research e.g at CMU [1], ICSI [2], FXPal [3], at NIST [4].

[5] describes the IDIAP smart meeting room environment and gives details about the hardware, on the web: `http://www.idiap.ch/~moore/meeting/`.

A first set of meetings have been captured called 'scripted meetings' that are 4-participant, 5-minute meetings with a given scenario . The work on action modeling [6] is based on this scripted meeting corpus that is now released publicly [7] . Another set of meetings has been recorded based on meeting data collection specifications described in [8] where there were less constraints on the scenario. More recently, the AMI meeting corpus has been recorded. More information about the corpus content is avaiable in The AMI Multimodal Meeting Database D2.2 deliverable. It is based on an elicitation task - design of a remote control - and some more meetings whose content is increasingly real ([9]). A part of the corpus data called hub set A was part of the 2005 NIST RT 05 evaluation.

The purpose of this document is to identify and describe the different steps performed to distribute recorded meetings.

Section 3 deals with the recording equipment setting before a meeting, including instructions to participants. Section 4 refers to the set of media (audio, video, slide captures, pen and whiteboard strokes) processing to run after a meeting have been recorded. Section 6 consists in the publishing, that is the secured web distribution of the data.

The different steps within each phase are described in chronological order, with some notes on what can be performed in parallel. Also we list current issues and what could be improved to maximize the automation of this process.

3 Smart Meeting Room Setup - Instructions: Before the Recording Media Acquisition

3.1 Video

The video signals are recorded on DV tapes with a timecode inserted on each video frame as shown in figure 1.

Four things have to be done related to DV tape preparation:

1. Prepare DVtapes with stickers
2. Put DVtapes in DVrecorders.
3. Check the input stream on miniDV by checking the video signals on DVrecorder screens. If necessary, unplug the FireWire cables on the two firewire hubs from Relax06 linux machine.
4. Check that all the DVtapes are rewinded: they should all start from the beginning.

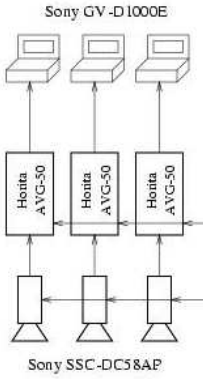

Figure 1: video acquisition.

3.2 Audio

Here are the steps for the configuration of the software that allows to record and synchronise the audio channels.

3.2.1 Creating a new template

1. Launch Samplitude software on idiap_smr windows machine. Create a new project with 22 audio channels or the number of mics you need (up to 22 because the first 2 channels of preamp C are not working).

   File -> New Virtual Project -> [NewName] . The name (meeting ID ie. IN1001) stands for directory name: J:\AMI\NewName .

   In Project Template select 22 mono track AMI.

2. Choose the input channel for each track. See the snapshot in figure 2.

3. Strict naming convention (Array1-1 to Array2-8, Lapel-1 to -4, Headset-1 to -4, Manikin-Left and -Right).

4. Save your project as a template.

3.2.2 Using an existing recording template

Here are five things to set so that the software correctly runs and saves the audio files.

1. Launch Samplitude on idiap_smr. Create a new project, select a folder and an existing template.

2. Click on "OK". A new folder will be created which will be the name of your project. All files will be stored there.

3. Right-click on a track (ie. lapel1), click on "track information", in record All tracks, select (ProjectName)_(TrackName).

4. Press backwards on the transport toolbar to set the timecode on midi timepiece at 00:00:00 while the one on the toolbar remains at 00:10:00.

5. Click on record the "Transport" toolbar to start recording from the selected microphones.

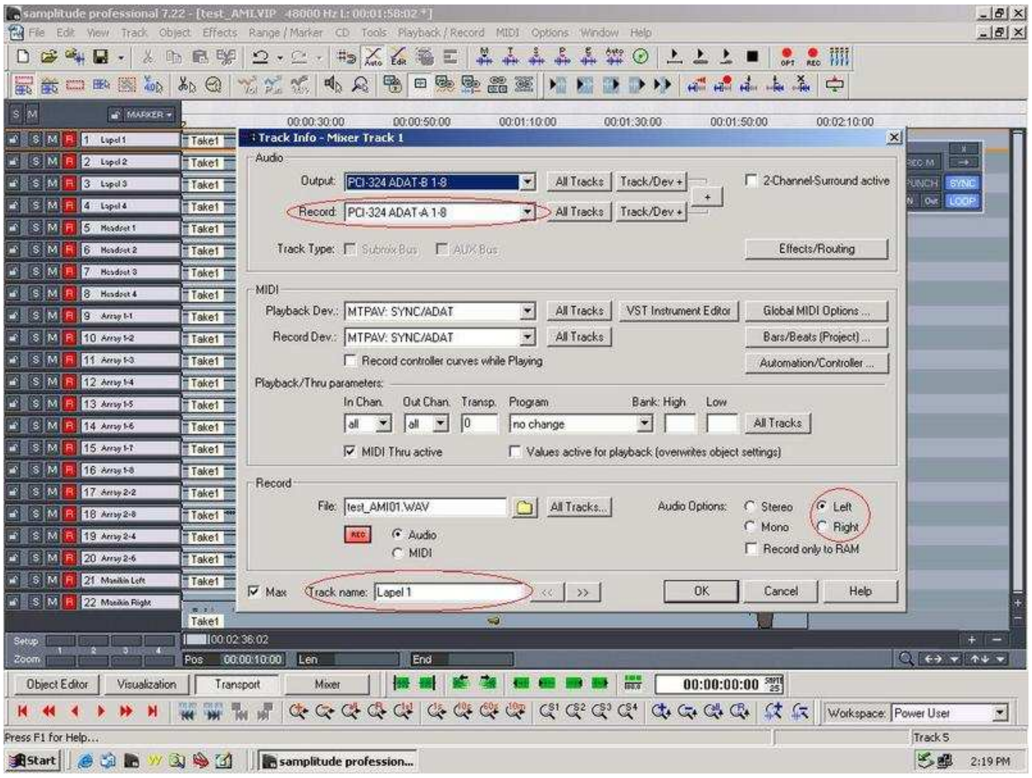

Figure 2: samplitude window: settings to change for a new template.

3.3 Whiteboard and slide capture

.The electronic whiteboard device and the projector from which the slide signal are captured are synchronised with the rest of the audio visual signals.

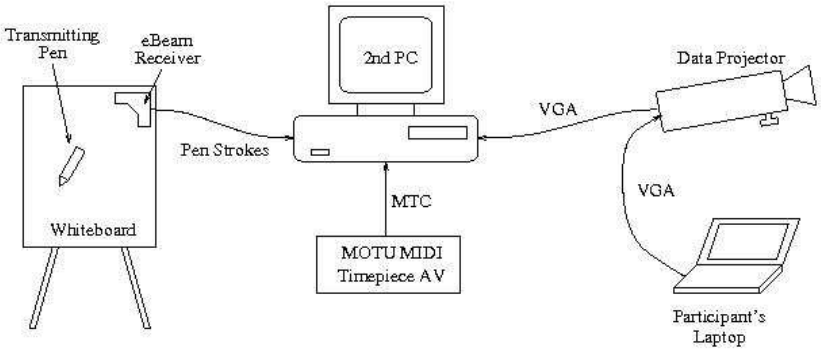

Figure 3: electronic whiteboard and slide projector set up.

1. Plug your laptop to the beamer. Select RGB or S-Video on the beamer.

2. Unplug and re-plug the USB cable on the e-Beam hardware device located on the whiteboard.

3. Launch e-Beam Sofware on Praline. The e-Beam device should be automatically detected. If not, unplug and re-plug the USB cable again.

4. Launch Meeting Capture (visible icon on desktop) on Praline, set the RGB capture settings (percentage of pixel change and time appearing on the screen for a new slide to be recorded, jpg picture quality), select a folder to record data in (the first windows asks you to browse to create an outputdir).

5. Right-click on the e-Beam Software icon in the task bar (bottom right of the screen), then left-click on "Disable system".

6. In the "Transport" toolbar of your samplitude project, press play if you want to record white-board and slides only. Press 'Record' to record also audio. Two windows should appear on Praline: one for the whiteboard and one for the slides.

7. Wait for the slide calibration to be done, then start your presentation. The slides will be stored as jpg files and whiteboard strokes as an xml file.

8. Quit Meeting Capture by clicking on "Exit".

3.4 Logitech pen

Each participant must take notes on the logitech notebook with their logitech pen. Synchronisation is done offline.

Figure 4: logiPen and logiBase

1. Before every meeting, download the files from all the pens.

2. Synchronize all the pens before every meeting by plugging them to the docking device. It only takes a few seconds.

3. Use the pens to write on the special paper during the meeting.

4. After every meeting, download the files from all the pens. Don't forget to rename the files with the pen number (ex Pen1_page012-08.03.2005.pen). * important *

5. Move the downloaded files to the folder you want them to be stored in.

6. After every meeting, don't forget to store the pen file creation time and the slide folder creation time in a text file. Other way there's no pen synchronization possible with other data.

3.5 Issues and proposed solutions to maximize automation

The main need of human intervention is due to hardware configuration and settings.

4 Recording

Here are the different things to check at the beginning, during and at the end of a recording.

Figure 5: SMR and SMR back

4.1 Instruction to participants

1. Handwritten notes should be performed on logitech paper only.

2. Hint for whitebaord: Be sure the whiteboard pens are buzzing when writing down or change the battery. Erase with special device.

4.2 Start recording

1. Launch Samplitude on idiap_smr windows machine. Create a new project using an existing template.

2. Click on record in the "Transport" toolbar.

3. When timecode starts, hit both REC and MARK buttons on the SONY remote control, check all video started, otherwise press both record buttons on DVrecorder itself and check Praline PC to make sure capture has started.

4.3 Issues at the start of a new recording

- Samplitude software issue:

  Though we set Samplitude time start at 10:00, and bring back the TimePiece timecode at 00:00, the audio starts recording from 00:00 instead of 10:00.

4.4 During Recording

1. Check that all videos have been started.

2. Check that there is enough space on firewire J:
disk linked to idiap_smr. The information is displayed in the record window in samplitude (ie. 28 hours of recordings). If there is not enough space, that might lead to audio dropout: the recording will stop.

4.5 Issues During Recordings

1. MidiTimePiece timecode stop issue: The clock on the MidiTimePiece hardware device on Idiap_smr windows machine stopped twice over 17 recordings.

4.6 Stop recording

1. Press Stop on Samplitude Record window: stops audio and slide/whiteboard capture.

2. Stop all videos with remote control by pressing STOP on the SONY remote control.

3. File -> Save project in Samplitude.

4. check capture has stopped too. Exit Meeting Capture window.

5. Put back correctly the headsets

4.7 Issues and proposed solutions to maximize automation

Ideally, the recording should start when people enter in the room, or by pressing one button.

5 Media processing

The idea behind this is to have a makefile that will call sequentially each of the following scripts as illustrated in figure 6.

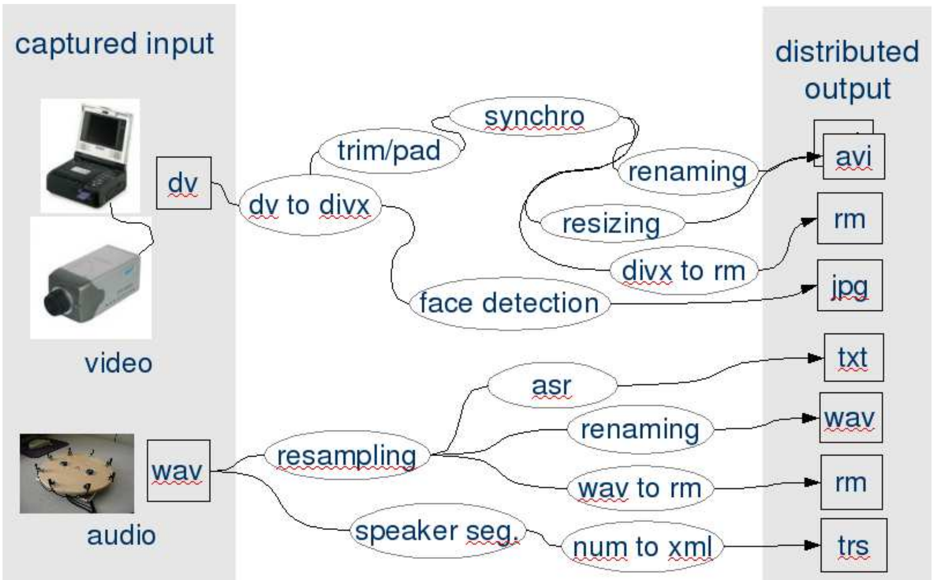

Figure 6: audio Video Processing steps.

5.1 Moving media files from local PC drive to MMM filer

At the end of the recording, the captured files should be moved from the local windows machine's disk to the mmm server for further processing.

1. The audio files are originally recorded at 48KBPS. They are all backed up on DVDs. A 16KBPS downsampling is performed for further processing. Copy audio 48K from external disk to MMM server.

   ```
   cp J:\AMI\<MeetingName> /mmm/shared/EXTRA/IDIAP_Meetings
   ```

   Create archive .rar for DVD backup.

2. Download logitech pen data. Put the 4 pens one after one on logitech pen base device to get the data.

3. Create a pens/ directory in the folder generated by Meeting Capture.

   mkdir D:\AMI\AMI\<MeetingName>\date_Time>\pens

4. The .pens files are automatically downloaded in:

   C:\Docs and Settings\smr\MyDocuments\MyioDocuments\<MeetingName>-penX-...pen

   Rename them following the filenaming convention

5. Copy them in the newly created pens/ directory:

   D:\AMI\AMI\<MeetingName>\date_Time>\pens

6. Copy slides/, pens/ and whiteboard/ on MMM server as well:

   mkdir /mmm/shared/IDIAP_Meetings/<MeetingName>

   copy D:\AMI\AMI\<MeetingName>\Date_Time in /mmm/shared/IDIAP_Meetings/<MeetingName>

5.2 Hierarchical structure creation

Data follows a strict naming convention. Meeting ID and filenames follow some rules explained in [10].

5.3 Captured input: Video

The goal of the video-processing step is to grab the video signals from DV tapes to DivX, repair the video signal where dropped frames occured so as to synchronise them and convert them in different format for streaming. This is show at the top of figure 6.

1. DV tape to DivX AVI conversion.

   - Create a new directory for the meeting on relax06, which will be called root directory.

     ```
     mkdir /scratch/IDIAP_Meetings/<MeetingName>
     ```

   - Copy the script /scratch/bin/multigrab_bastien to the root dir. Modify the script by 1) selecting the cams from which the videos will be grabbed (put in comments or not), 2) setting video duration.
   - Connect the cams, rewind the tapes. Notice that only 3 cams can be plugged on each hub at the same time, so you'll have to do the same job again for the 7th one.

```
##### VAR ##########################
TIMER=00:45:15:00
TIMEC=00:45:14:00
TIMEL=01:45:16:00
TIME1=00:45:14:00
TIME2=00:45:18:00
TIME3=00:45:15:00
TIME4=00:45:19:00

$DVGRAB --guid $IDR --duration smpte=$TIMER | $ENCODER - -demuxer 22 -cache 1024
        -nosound -ovc lavc -lavcopts vcodec=mpeg4:vhq:vbitrate=2300:keyint=25
        -o $FILE_R.avi
rm dvgrab*.avi
$CORRECT $FILE_R.avi
```

- Check that relax06 /scratch must have one third of free space otherwise there might have encoding problems.
- Launch the script ./multigrab_bastien from the root dir and wait for the data digitalization to end.
- Remove and archive tapes in box
- Move the newly created files to the corresponding directories (ex: mv *.avi video/)

2. Video renaming
   - Video rename: /scratch/tools/renameIDIAP.pl: modify and launch
   - Create a directory in /com/mmm/shared (G:\shared with windows). This new directory should have the same name as the root dir.
   - Copy the local data to newly created network directory.

3. Video synchronisation
   - Log file generation: A log file is a text file containing a list of all frame drops and timecode issues for a given video. We use it to repair the videos by inserting frame(s) where some are missing.
     create log files: checks missing frames wrt timecode.

```
foreach AVIvideo (‘ ls /.../*avi‘)
  /mmm/tools/tempDebug/AviControl -log  -verbose $AVIvideo
end
```

   - Checking log files and correct if necessary. A first pass consists of updating automatically the log file (checking). In case of weird and messy log file, a 2nd human update is performed.

```
foreach i (‘ ls /.../*avi‘)
mkdir /.../LOG-ORIG
cp /.../*log /.../LOG-ORIG
```

```
-> write a *log.temp in the same dir as the log when the log is
   correctly formated.

foreach logFile (‘ ls /.../*log‘)
  ~/Mael/ReRunaviReair/formatLogV2 log=$logFile
end
```

- Repair video in case of dropped frames.
  Modify /mmm/tools/ReRunAviRepair/IN1008-9.pl

```
foreach avi (‘ ls /.../*avi‘)
   /mmm/tools/ReRunAviRepair/IN1008-9.pl video=$avi
 end
```

Here is an example of Log file:

```
AviControl CmdLine
Avi file: IB4011-R_T000023.680_T004011.760.avi
Log file: IB4011-R_T000023.680_T004011.760.log
[mpeg4]  720x576  60294 frames  25 fps
Total Encoding Errors:     10
Percentage Encoding Errors:   0

<<< AVI file chunked into 20 pieces(s) :>>>
  TimeCode Begin           :  00:00:23:17 (frame:      0)
  TimeCode Correct Run     :  00:00:23:17 (frame:      0)  ->  00:08:59:08 (frame:  12891)
  TimeCode Frame Drop      :  00:08:59:08 (frame:  12891)  ->  00:08:59:10 (frame:  12892)
  TimeCode Correct Run     :  00:08:59:10 (frame:  12892)  ->  00:10:33:21 (frame:  15253)
  TimeCode Frame Drop      :  00:10:33:21 (frame:  15253)  ->  00:10:34:00 (frame:  15254)
  TimeCode Correct Run     :  00:10:34:00 (frame:  15254)  ->  00:10:48:22 (frame:  15626)
  TimeCode Frame Drop      :  00:10:48:22 (frame:  15626)  ->  00:10:49:13 (frame:  15627)
  TimeCode Correct Run     :  00:10:49:13 (frame:  15627)  ->  00:11:49:01 (frame:  17115)
  TimeCode Frame Drop      :  00:11:49:01 (frame:  17115)  ->  00:11:49:08 (frame:  17116)
  TimeCode Correct Run     :  00:11:49:08 (frame:  17116)  ->  00:11:49:09 (frame:  17117)
  TimeCode Frame Drop      :  00:11:49:09 (frame:  17117)  ->  00:11:49:20 (frame:  17118)
  TimeCode Correct Run     :  00:11:49:20 (frame:  17118)  ->  00:11:59:02 (frame:  17350)
  TimeCode Frame Drop      :  00:11:59:02 (frame:  17350)  ->  00:11:59:20 (frame:  17351)
  TimeCode Correct Run     :  00:11:59:20 (frame:  17351)  ->  00:12:39:06 (frame:  18337)
  TimeCode Stop            :  00:12:39:06 (frame:  18337)  ->  00:12:39:06 (frame:  18338)
  TimeCode Frame Drop      :  00:12:39:06 (frame:  18338)  ->  00:12:39:08 (frame:  18339)
  TimeCode Correct Run     :  00:12:39:08 (frame:  18339)  ->  00:17:15:24 (frame:  25255)
  TimeCode Frame Drop      :  00:17:15:24 (frame:  25255)  ->  00:17:16:02 (frame:  25256)
  TimeCode Correct Run     :  00:17:16:02 (frame:  25256)  ->  00:26:12:07 (frame:  38661)
  TimeCode Frame Drop      :  00:26:12:07 (frame:  38661)  ->  00:26:12:10 (frame:  38662)
  TimeCode Correct Run     :  00:26:12:10 (frame:  38662)  ->  00:40:37:16 (frame:  60293)
  TimeCode End             :  00:40:37:16 (frame:  60293)
```

Timecode stop appears when the log generator reads a same timecode on consecutive frames.

4. Trim / Pad

   - AVS generator
     *.avs files are used by VirtualDub for trimming or padding videos to a given start time. This helps audio and videos to be perfectly synchronised.
     Modify /Mael/ReRunAVIrepair/generateIS-AVS-oct18.pl

     ```
     avisource("K:\secret\ES10-16\REPAIRED\IB4002-Closeup2_T000006.880_T003215.080.avi")
     Trim(78,0)
     ```

     This file aims at cutting the first 78 frames corresponding to 10s - 6.880s (available form the video filename) = 3.120 seconds.

5. Multiple video encoding

   - Generate FFMPEG resized and cropped videos
     Small size videos are used by annotators with NXT, Event Editor or other annotation tools in order to view multiple videos without rendering probems.

     ```
     /com/mmm/tools/ffmpeg -i
     inputAVI.avi
     -s cif -f avi -an -vcodec mpeg4 -b 800
     outputAVI.avi
     ```

   - Test audio-video, video-video synchronization with Event Editor on windows.
   - DivX to RM conversion
     Generating makeRM.bat by modifying and launching:
      /Bastien/MakeRM/createRMjob.pl
   - Launch the .bat file from Praline or another windows PC with RealProducer. This will convert avi videos to realmedia rm. Realmedia is used for audio video streaming.

6. Face detetion

   Generating JPG face images from closeups using Sebastien Marcel's face detector detailed in `http://www.idiap.ch/~marcel/en/torch3/torch3vision.php`.

   ```
   example in ~/Mael/TNO/TNOfaces/detectTS3000.pl

   cd /home/mmmdev/Mael/NewFaceDetector

   ./facedetect-video $path -verbose -vitc -I 46000 -threshold 0.1 -F
   -crop -minsize 30 -maxsize 800 -drop 25 -o $videoDir/closeup -jpg

   + jpg rename
   ```

5.4 videoIssues

- Frame Drop Issues: In some cases, the log file derived from the video includes a lot of frame drops, reset, stops and is very difficult to interprete: too many different cases to handle automatically.

  ```
  example on /com/mmm/...
  ```

5.5 audio

1. Downsample the audio files from 48kHz to 16kHz by using GoldWave. An alternative is sox but resulting quality is not as good:

   ```
   sox in.wav -r 16000 out.wav
   ```

   - Launch GoldWave
   - File -> Batch processing
   - Add folder : select the folder in wich 48k audio files are stored.
   - Select 16kHz downsampling and the options 'store in the same folder' and 'overwrite existing files'

   Copy the newly created 16kHz audio files in /mmm/shared/IDIAP_Meetings/<MeetingName>/audio/

2. Audio mixing for headsets and lapels

   ```
   cd /com/mmm/tools/
   ./my-mix-it *lapels*wav lapelmix.wav
   ```

3. Audio wav to realmedia conversion

   ```
   /usr/local/realnetworks/producer10/producer -i inputFile.wav -o
   outputFileRealMedia.rm -a IDIAP -c IDIAP -dv -dt -ad "16k rv8" -am
   voice
   ```

5.6 Speaker segmentation and Transcriber formating

1. matlab software from Guillaume

   ```
   example in /home/vision/guillemo/SMR-SEG/MULTISEG/mySegES2009.m
   ```

2. awk script to convert from *.txt to *.trs (xml):

   ```
   awk -v
   outputTRSfile=preTranscript.trs
   -f /home/mmmdev/Mael/TNOdata/speakerSeg2XMLnew.awk seg.txt
   ```

3. txt to multiple xml in the browsable directory

```
example in ~/Mael/TNO/TNObrowsable/script-AMI-IDIAP-MeetingsBrowsable.pl

system("awk -v output=$sessionDir/browsable/$meetingDir-speakersegAll.xml
        -f spksegToSegAll-TNO.awk $sessionDir/other/*.txt");

system("xsltproc -o $sessionDir/browsable/$meetingDir-speakersegChan0.xml
        copychan0.xsl $sessionDir/browsable/$meetingDir-speakersegAll.xml");
system("xsltproc -o $sessionDir/browsable/$meetingDir-speakersegChan1.xml
        copychan1.xsl $sessionDir/browsable/$meetingDir-speakersegAll.xml");
system("xsltproc -o $sessionDir/browsable/$meetingDir-speakersegChan2.xml
        copychan2.xsl $sessionDir/browsable/$meetingDir-speakersegAll.xml");

system("xsltproc -o $sessionDir/browsable/$meetingDir-speakersegChan3.xml
        copychan3.xsl $sessionDir/browsable/$meetingDir-speakersegAll.xml");
```

The speaker segments in xml format and the transcript in HTML go into the browsable directory.

5.7 info.xml and Posgres Database

- structure of the info.xml

  This xml file contains the static metadata information of the meeting such as the participant names, the date, the time and the type of the meeting. It is mainly used by Ferret meeting browser. An example of such a file is the following where the duration is set in seconds. In AMI scenario-based meetings, each person has a role [PM Project Manager, ME Marketing Expert, UI User Interface Designer, ID Industrial Designer] and a name. The id attribute corresponds to a closeup camera or a seat where a participant is at the beginning of a recording.

```
<?xml version="1.0" encoding="UTF-8"?>
<meeting

meetingId="IS1008c"
name="AMI scenario meeting"
dateOnly="2004-12-10"
startTime="13:59:00"
description="AMI scenario Hub - meeting"
duration="1546"

>

<person id="4" name="FIE073" role="ID"/>
<person id="2" name="FIE038" role="UI"/>
<person id="1" name="MIO086" role="PM"/>
<person id="3" name="MIE085" role="ME"/>

</meeting>
```

  Role attribute is only provided with AMI scenario meetings.

- Posgres Database on relax17 http://relax17.idiap.ch/phpmyadmin/

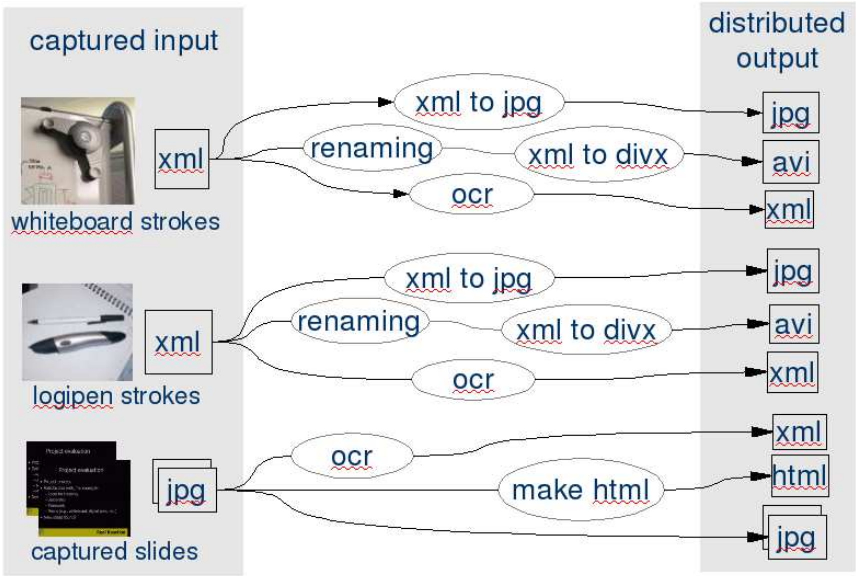

Figure 7: Slide, whiteboard and pen strokes processes.

5.8 slides

script template in /Mael/TNO/TNOslides . Optical Character Recognition ([11]) is performed on each detected slides.

1. renameJPGslides-TS3000-1-2.pl renames the JPG slides.
2. makeHTMLslides.pl creates HTML pages for browsing: index, left, slides.html.
3. OCR: examples in /home/mmmdev/Mael/OCR/

5.9 logitech pen data and whiteboard

The process for converting the xml whiteboard strokes and logitech output is based on Marcus liwicki's software ([12]).

1. Examples for generating videos (For the new Parameters see below):

```
Whiteboard.exe -p -f IS1003d-strokes.xml IS1004d-strokes.xml

//Creates two videos (one for each xml-file) with past- and future shadow

Whiteboard.exe -f -p -s1 -b3 -m2 IS1000d-pen_3.pen IS1000d-pen_4.pen

//Creates two videos (one for each pen-file) If a pen-File is not in
LogiPen1-format the program will skip this file and it will print
out: "no stoke data"
```

2. Issues.. Logitech pen strokes are not synchronised with the rest.
3. Logitech pen data is sometimes not uploaded properly.

4. Logitech pen files and whiteboard XML strokes are converted on windows OS due to the use of Microsoft Visual C++ library.

5.10 Issues and proposed solutions to maximize automation

All this set of processes should ideally run without human intervention.

6 MMM Web Distribution

6.1 Web access security and user authentication

Part of the data on the MMM server is accessible only for IM2 and AMI project partners.

Several tools are used to constrain access via a username and password authentication.

1. .htaccess in /com/mmm/data/private/AMIzone/
2. filemanager tool: /com/mmm/protected/

6.2 Data on MMM server

The data on MMM server is organized in multiple places:

1. The M4 scripted meeting corpus in /mmm/data/M4-Corpus/
2. Use of symbolic links and randomly generated directories make the AMI corpus to be protected.
3. automatic generation of directories and sym links:
4. Media Structure

   audio/ browsable/ info.xml other/ pens/ slides/ video/ whiteboard/
5. Annotations

   Annotations covers all modalities including ortographic transcription, dialogue acts, abstractve summaries, named entities, head and hand gesture, gaze etc.

6.3 MMM services

Mainly Perl CGI (Common Gateway Interface) and JSP (Java Servlet Pages) are used as dynamic web page generators. Once the data is correctly organized in directories following the naming convention, the role of Perl CGI / JSP is to look for the files in the server and publish them on the web.

1. Perl CGI scripts:

   ```
   /com/mmm/tools/cgi-bin/amiHub/
   ```
2. Java Servlet Pages:

   ```
   /com/mmm/tools/tomcat/webapps/ROOT/amiHub/
   ```
3. Ferret meeting browsing interface software:

   ```
   /com/mmm/tools/tomcat/webapps/ROOT/ferretAMI/
   ```

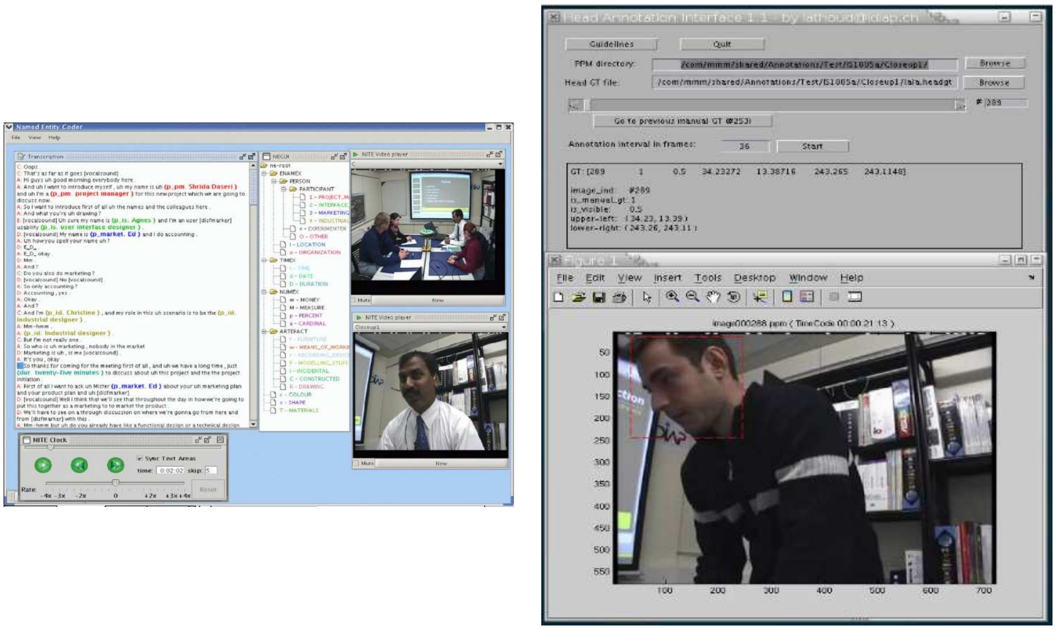

Figure 8: NITE XML Annotation Toolkit (NXT) and a Matlab Head Annotation GUI

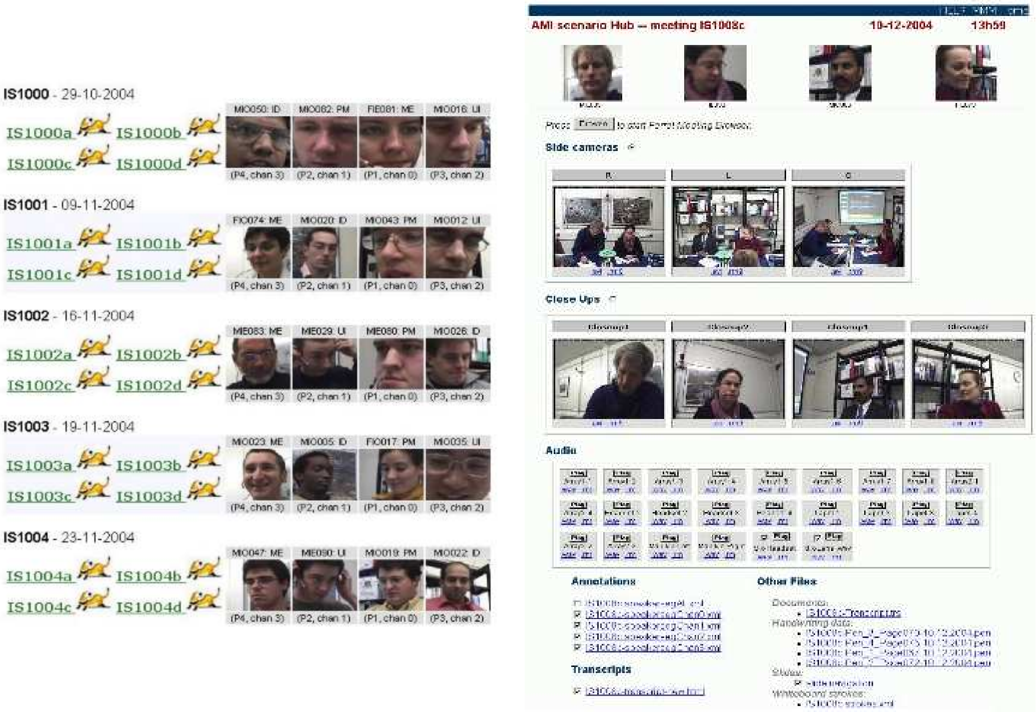

Figure 9: web GUIs to (a) select a particular meeting and (b) to select media and annotations in a particular meeting.

6.4 Backups

Media are backed up on DVDs: 48K audio , and media after having processed are saved on DVDs: videos in AVI, RM and with a log file, audio in 16K WAV and RM, slides in JPEG, whiteboard and

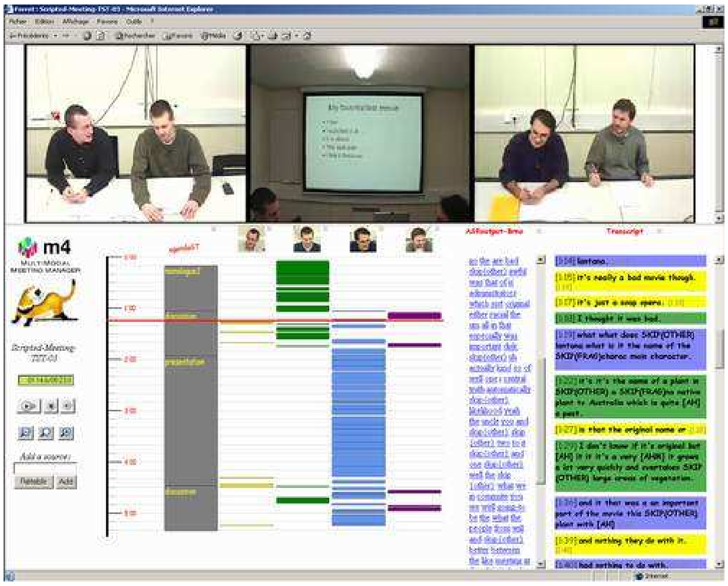

Figure 10: Ferret meeting browser.

pen strokes. We also keep original DV tapes Regular snapshots / backups on hard-disk (rsynch / synchro). For details about media file server hardware, check [13].

7 Conclusion

The goal of this report was to list the set of processes to be applied to meeting recordings. Also we highlighted the different issues in making an ideal Makefile doing everything automatically.

To sum up the issues, the main one is the fact that different machines and OS are needed to do the processing. Also, videos are recorded as DV tapes which needs human intervention for further conversion. And the video signal usually has a couple of missing frames due to the grabbing step that is repaired semi-automatically.

References

[1] Susanne Burger, Victoria MacLaren, and Hua Yu. The isl meeting corpus: The impact of meeting type on speech style. In *Proceedings of the ICSLP '02*, September 2002.

[2] Adam Janin, Don Baron, Jane Edwards, Dan Ellis, David Gelbart, Nelson Morgan, Barbara Peskin, Thilo Pfau, Elizabeth Shriberg, Andreas Stolcke, and Chuck Wooters. The icsi meeting corpus. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 364–367, 2003.

[3] Patrick Chiu, Ashutosh Kapuskar, Lynn Wilcox, and Sarah Reitmeier. Meeting capture in a media enriched conference room. In *Cooperative Buildings*, pages 79–88, 1999.

[4] Nist smart meeting room.

[5] D. Moore. The idiap smart meeting room. IDIAP-COM 07, IDIAP, 2002.

[6] I. McCowan, S. Bengio, D. Gatica-Perez, G. Lathoud, F. Monay, D. Moore, P. Wellner, and H. Bourlard. Modeling human interaction in meetings. In *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, April 2003.

[7] M4 corpus.

[8] I. McCowan, D. Gatica-Perez, and S. Bengio. Meeting data collection specifications. IDIAP-COM 10, IDIAP, 2003.

[9] J Carletta, S Ashby, S. Bourban, M Flynn, M Guillemot, T Hain, J Kadlec, V Karasaiskos, W Kraaij, G Lathoud, M Lincoln, A Lisowska, I McCowan, W Post, D Reidsma, and P Wellner. The ami meeting corpus: A pre-announcement. In *Workshop on Multimodal Interaction and Related Machine Learning Algorithms (MLMI)*, 2005.

[10] Ami multimodal meeting database naming convention.

[11] Datong Chen, Jean-Marc Odobez, and Jean-Philippe Thiran. Monte Carlo Video Text Segmentation. *accepted for publication in Int. Journal Of Pattern Recognition and Artificial Intelligence*, 2005.

[12] M. Liwicki and H H. Bunke. Enhancing training data for handwriting recognition of whiteboard notes with samples from a different database. In *Proceedings of the 8th Int. Conf. on Document Analysis and Recognition*, 2005.

[13] Frank Formaz and Norbert Crettol. The idiap multimedia file server. IDIAP-Com 05, IDIAP, 2004.