# 1. ICSI Meeting Setting

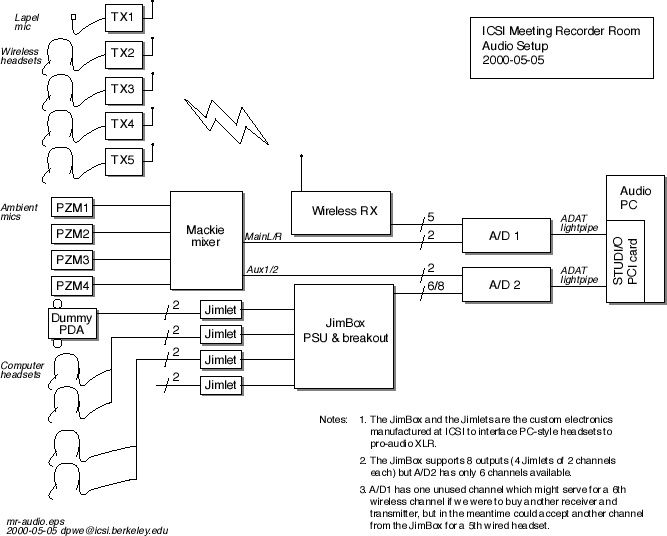
<http://www1.icsi.berkeley.edu/Speech/mr/mtgrcdr.html>

## Conference Room



## Setup

The equipment is installed in conference room 6A, which has a small machine room at the rear. Most of the equipment is in the machine room, with the cables from the microphones coming down from the conference table and through a hole in the intervening wall. The overall block diagram is shown below:



### The equipment that have used is:

* Wireless microphones: [Sony ECM-310BMP headsets](http://bpgprod.sel.sony.com/model.bpg?cat=Pro+Audio&subcat=Microphones%2c+Wireless&model=ECM310BMP) and a Sony ECM-77BMP lapel mic feeding into [Sony WRT-805A body pack transmitters](http://bpgprod.sel.sony.com/model.bpg?cat=Pro+Audio&subcat=Microphones%2c+Wireless&model=WRT805A).
* Wireless receiver: [Sony MB-806A modular base unit](http://bpgprod.sel.sony.com/model.bpg?cat=Pro+Audio&subcat=Microphones%2c+Wireless&model=MB806A) containing five [Sony WRU-806A/64 UHF synthesized tuner modules](http://bpgprod.sel.sony.com/model.bpg?cat=Pro+Audio&subcat=Microphones%2c+Wireless&model=WRU806A68).
* Analog-to-digital converters: Two [Sonorus AUDI/O AD/24](http://www.sonorus.com/audio/) 8 channel 24 bit 44/48kHz optical-output converters, connected to a [Sonorus STUDI/O 16 channel digital audio PCI bus interface](http://www.sonorus.com/studio.html), mounted in a Dell 500 MHz Pentium III PC running Red Hat Linux 6.1.
* Ambient mics: Four [Crown PZM-6D](http://www.crownaudio.com/crownaudio/mic_htm/pzm.htm) pressure zone microphones.
* Home-made dummy PDA with two cheap electret elements mounted about 8cm apart.
* In-house-manufactured bias power/amplifier/line driver to interface conventional computer headsets with pro-audio XLR inputs.

### Recording software

Computer

The meeting room audio computer, popcorn.aciri.org, is a Dell 500 MHz Pentium III with 256MB of RAM and an internal 10GB SCSI-2 disk. It is running RedHat Linux 6.1, with a 2.2.12SMP kernel.

The 16 channel audio comes in over 2 ADAT lightpipe optical feeds into a Sonorus STUDI/O PCI interface card. Sonorus have given us a beta version of a driver for that card that works within the [ALSA sound driver architecture](http://www.alsa-project.org/). I have made some slight customizations to the version of the driver (snd\_studio.o) that is running on that machine, but I expect my software will be compatible with future versions of the driver. The source of the driver is currently on the internal disk under /u/dpwe/projects/sonorus/studio-driver-0.3.3a/; this directory is a copy of the directory with the same path on the main ICSI file system, and the two are supposed to be kept in sync, although it is a manual process.

### Software

The recording software is composed of two pieces:

**rcddrv**, a driver program that handles reading data from the ALSA driver and writing it to files on disk, and

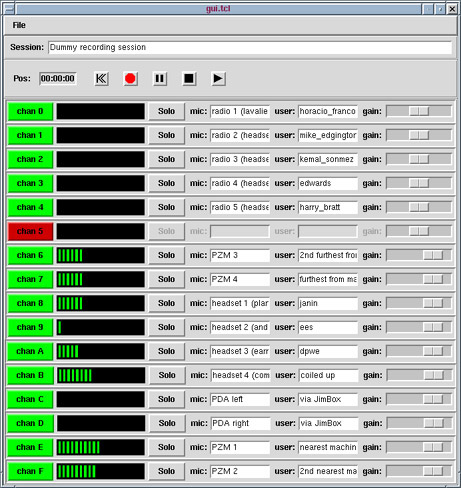
**rcdgui**, a Tcl/Tk graphical user interface that provides a glossy way to control rcddrv, as well as managing the ancilliary information describing the recordings that is usually stored in the KEY file (channel identities, recording date etc.).

**rcddrv** is a C program based on the example code supplied with the Sonorus driver. Its source is currently in /u/dpwe/projects/sonorus/studio-driver-0.0.3a/test/ (although I should rename that to rcddrv soon). It accepts a simple command language via standard input, including some information reports to standard output. This language is not documented except within the source code and the way it is used by rcdgui. One possible advantage of having rcddrv as a separate program is that it could be run at a higher priority, to reduce the risk of dropped samples. However, this is seemingly not required at present.

**rcdgui** runs under the wish Tcl/Tk interpreter, although it also loads the [incr Tcl] object-oriented extensions. A screenshot is shown below. Each input channel has a separate 'mixer bar' element that allows activation/deactivation of that channel (to avoid the overhead and disk use of writing channels that are not being used), gain adjustment (which is actually the number of bits, between 0 and 16, by which the samples are shifted up before being stored; only the top 16 bits are saved to disk, so some gain adjustment may be required to optimize the dynamic range), and text entry widgets for the microphone ID and the associated participant, if any.

The mixer bar also includes a VU-style meter for each channel, based on the peak levels which are tracked by rcddrv. These levels reflect the influence of the gain slider to aid in level setting. Even inactive channels (indicated by the dull red channel button and greyed-out entry boxes) continue to show peak level displays, to alert the user if a channel that is believed to be disconnected actually has a signal coming in.

Note: the SOLO button has no effect at present, and neither do the play or return-to-zero buttons of the transport. Further, although pause and stop pass different commands to rcddrv, they have the same effect - you can resume recording after stopping, and the new data is simply appended to what has been recorded already. The only way to discard your recording to date is to quit the program and start over.

[](http://www1.icsi.berkeley.edu/~dpwe/research/mtgrcdr/rcdgui.gif)

The session name, microphone and user identities, gains and channel activations can be saved to a "KEY" file under the File menu. An existing KEY file can also be read back, restoring these attributes. KEY files also contain fields for Description, Participants and Notes, each of which can be edited in a pop-up window accessed through the File menu. The date and time of the recording are automatically recorded in the KEY file.

The source to rcdgui is currently in /u/dpwe/projects/sonorus/studio-driver-0.0.3a/gui/, which should become rcdgui.

# 2. AMI Meeting Setting

<http://groups.inf.ed.ac.uk/ami/corpus/>

## Audio setup

The room contains 24 microphones from which 24 mono audio channels are recorded directly to hard disk. 16 Sennheiser MK2E-P-C miniature omni-directional electret microphones are arranged in two 10cm radius circular arrays of eight. These are placed in the center of the meeting room table, one between the participants and one at the end of the table closest to the presentation screen and whiteboard. The MK2E-P-C was chosen for its 20Hz-20kHz linear frequency response and its ability to draw phantom power directly from themicrophone pre-amplifier. Eight Sennheiser EW300 Series radio microphones are used for recording the four meeting participants. Each person wears an ME 3-N close talking headset condenser mic and an MKE 2-EW omni directional lapel mic -- the wireless equivalent of those used in the microphone arrays. Use of a radio based system allows participants to move freely around the room without diminishing the quality of audio recordings.

Three Focusrite Octopre eight-channel microphone pre-amplifiers with up to 24bit 96kHz analogue-to-digital converters are used to amplify and digitize the microphone outputs. Each channel has a separate class A amplifier with independent gain control. Digitized output is via a single ADAT Lightpipe fiber optic cable carrying all eight channels. The A-to-D converters can sample at a variety of rates either using the Octopre's internal clock or from an external source via a word-clock input. Here the data is captured at a 48kHz, 16bit resolution. The Octopres also provide phantom power for the MK2E-P-C microphones.

The Mark of the Unicorn (MOTU) 2408 MKIII is an audio interface for PC based hard disk audio recording. It consists of a 19-inch rack-mounted I/O unit connected via a Firewire-like interface to a PCI card. The I/O unit supports 24 input/output channels in three banks of eight, with all 24 channels capable of operating simultaneously. Software installed on the PC allows configuration and acquisition of each of the channels via the PCI card. In the meeting room, each of the ADAT Lightpipe outputs from the Octopre A-to-D converters are connected to one bank of a single I/O unit and are subsequently acquired by the PC via the PCI card.

The audio capture computer is a 3GHz P4 with two 40MB SCSI hard drives configured as a RAID 0 array for streaming audio. The operating system used is Windows XP for compatibility with the MOTU driver software. Audio is captured and exported using Cakewalk Sonar recording software.

## Video setup

Six cameras are used to record video proceedings. Four [Sony XC555 subminiature cameras](https://pro.sony.com/bbsc/ssr/cat-camerasindustrial/cat-ciindustrial/product-XC555/) with 6mm lenses mounted under the central microphone array provide close-up views of each of the meeting participants. Two Sony SSC-DC58AP CCTV cameras, each with a 3.6mm semi-fisheye lense, provide wide-angle views of the room. One is mounted above the center of the table and gives an overhead view of the entire floor area of the room. The other is mounted in the corner of the room and provides a view of the whiteboard and presentation areas.

Six Sony GV-D1000E digital video recorders are used to record the output of the cameras directly to Mini-DV cassettes. Using Mini-DV provides reliable video capture with few errors or dropped frames. It also provides an immediate tape backup of the raw video data.

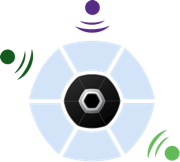
Synchronization  
Special hardware is used to provide synchronization signals. Global time-stamping allows the A-to-D converters in the Octopres to sample each channel at the same time, thereby avoiding a time skew between audio channels. Cameras also acquire frames at the same time, avoiding lags between video channels. The Horita BSG-50 PAL Blackburst Generator generates a composite video timing signal which is used as a reference signal to which all other devices are locked. The signal is fed directly to each of the video cameras to ensure that they sample frames at the exact same time. A further output is connected to a MOTU MIDI Timepiece AV, which generates all other timing signals. The MOTU MIDI timepiece AV (MTP-AV) is capable of locking to and generating a number of different timing signals. In the meeting room, the MTP-AV locks to the Blackburst reference signal and generates a 48kHz word clock for triggering the A-to-D converters in the 3 Octopres. This ensures that each audio channel is sampled at precisely the same instant. The MTP-AV also creates a Longitudinal Time Code (LTC) for each video frame. The LTC is encoded as an 80-bit word (Hours:Minutes:Seconds:Frames) and output as a 2kHz audio signal. In addition, the MTP-AV outputs a MIDI Time Code. This is the LTC output in a format which can be read by MIDI devices. In the meeting room, it is read by the Sonar recording software and used to time-stamp the audio samples.

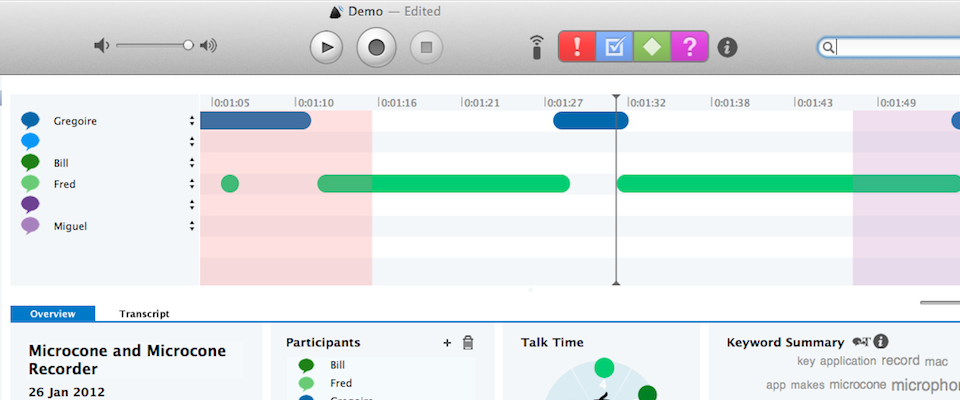
The Horita AVG-50 time-code inserters translate the 80bit LTC audio signal into a 90bit Vertical Interval Time Code. This 90-bit code is then inserted into the top two lines of each video frame as a series of black and white blocks, which may subsequently be read during video playback. Since this code corresponds directly to the Midi Time Code being used to time stamp the audio recording, precise synchronization of the audio and video signals can be achieved.

# 3. Other Solution

## [Microcone](http://www.dev-audio.com/products/recorder/)

* Single microphone array device up to 6 channels
  + pros
    - Easy to set up (one single device)
    - Both software and hardware are supported
  + Cons
    - Have to be placed into the center



## [Voice TracerMeeting recorder](https://www.dictation.philips.com/us/products/product/voice_tracer_meeting_recorder_dvt7000/overview/)



* Like Microcone (3 mic)
  + Pros
    - wireless remote control

## VS4Recorder

It is a single PCIe slot which enables four independent HD inputs with up to eight embedded audio channels per source.

* Pros
  + Provide software to synchronize the Video and Audio
* Cons
  + In my understanding, the audio comes with the video

## Individual Device

### Audio

Clear One (<http://www.clearone.com/>):

* [Tabletop Microphone](http://www.clearone.com/products_tabletop_mic) (Used by Jeffrey)



### Camera

Point Grey (<http://ww2.ptgrey.com/>)

* It offers high quality camera.

# References:

Janin, Adam, et al. "The ICSI meeting corpus." *Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on*. Vol. 1. IEEE, 2003.

McCowan, Iain, et al. "The AMI meeting corpus." *Proceedings of the 5th International Conference on Methods and Techniques in Behavioral Research*. Vol. 88. 2005.