Real Time Multi-Modal Communication: Audio Interface

Version 0.2 - August, 1999

Peter Søren Kirk Hansen,

and

John Aasted Sørensen

Department of Mathematical Modelling
Section for Digital Signal Processing
Technical University of Denmark
Building 321, DK-2800 Lyngby, Denmark

IMM

Chapter 2

Audio Specifications

The specifications of the audio agent in the RTMM video conferencing system are based on the desired functionality, and results in a number of requirements to the system necessary to achieve this functionality. From the system layout, the technology/algorithms used to implement the different subtasks can be considered, together with the interface to other agents in the video conferencing system. Finally, the hardware components can be chosen, and the physical setup of the system can be discussed.

2.1 Functionality

The preliminary functionality of the audio part of the RTMM video conferencing system has been defined by the Department of Information and Media Science at the University of Aarhus. The desired functionality is:

- High quality sound with clear intelligibility and speaker identification. Dropouts can not be tolerated.
- Suppression of especially impulse noise from papers, cups, etc.
- Simultaneously sound from all units in a multi-connection setup.
- The listeners should be able to identify the speakers in a virtual seminar room. This gives the user a more natural communication impression, makes it easier to follow a discussion and therefore, increases the users acceptance.
- Intelligent speaker tracking can be used for camera and video control.
- Both special video conferencing rooms and desktop environments should be supported, i.e., different network and platform requirements.

2.2 Requirements

The most important requirement for the audio part of the RTMM video conferencing system is high quality. This is expected to be achieved by careful design of the critical pre-processing steps:

- Multi-microphone based sound capture is necessary for several pre-processing steps. The solution should be insensitive to microphone characteristics and spatial setup.

- The natural sound perception depends on reverberations and can only be obtained by beamforming. The algorithm should be adaptive with slow convergence and wide main loop.
- Stereo echo cancellation is necessary in order not to have acoustic feedback from loudspeakers to microphones.
- Suppression of impulse and annoying background noise.

The coding of audio streams must be bandwidth efficient and flexible, i.e., layered encoding for heterogeneous receivers and activity detection:

- Encoding algorithm that produces multiple layers, or quality, such that sites with different bandwidth connections can communicate.
- Voice/sound activity detection can be used to improve transmission and decoding efficiency. Thus, sites with no activity only transmit information about the background noise, which is used to generate comfort noise at the receiving sites, i.e., typically only one decoder corresponding to the speaker active site is running.
- There must be a separate decoder or comfort noise generator for each of the other sites.
- Low encoding/decoding delay is important (less than the video based delay).
- The decoder must implement strategies to recover from package loss, such that graceful degradation is assured.
- Timing information (system clock) must be available in order to synchronize audio with other information in the session.

The post-processing steps are used to make the virtual seminar room comfortable as if everyone was together. This includes:

- Floor control which can be used in one-way communication situations, and to mute disturbing sites.
- Site tracking used for video control, i.e., showing the speaker active site.
- The virtual seminar room is created by combining audio signals from all units into a stereo signal such that the sound from each site comes from different spatial directions.

Speaker tracking used for camera control and speaker identification used for video control are desired functionality, which is out of scope in this project. Thus, the considered requirements for the audio part results in a layout as shown in Figure 2.1.

2.3 Technology

The technology used to implement the different requirements is only sporadic defined:

- The analog audio signals are sampled with 48 kHz, using 24 bit resolution.
- The MPEG standard is suggested for encoding of the audio signal. However, maybe it is necessary to use CELP coding for low bit-rate links.

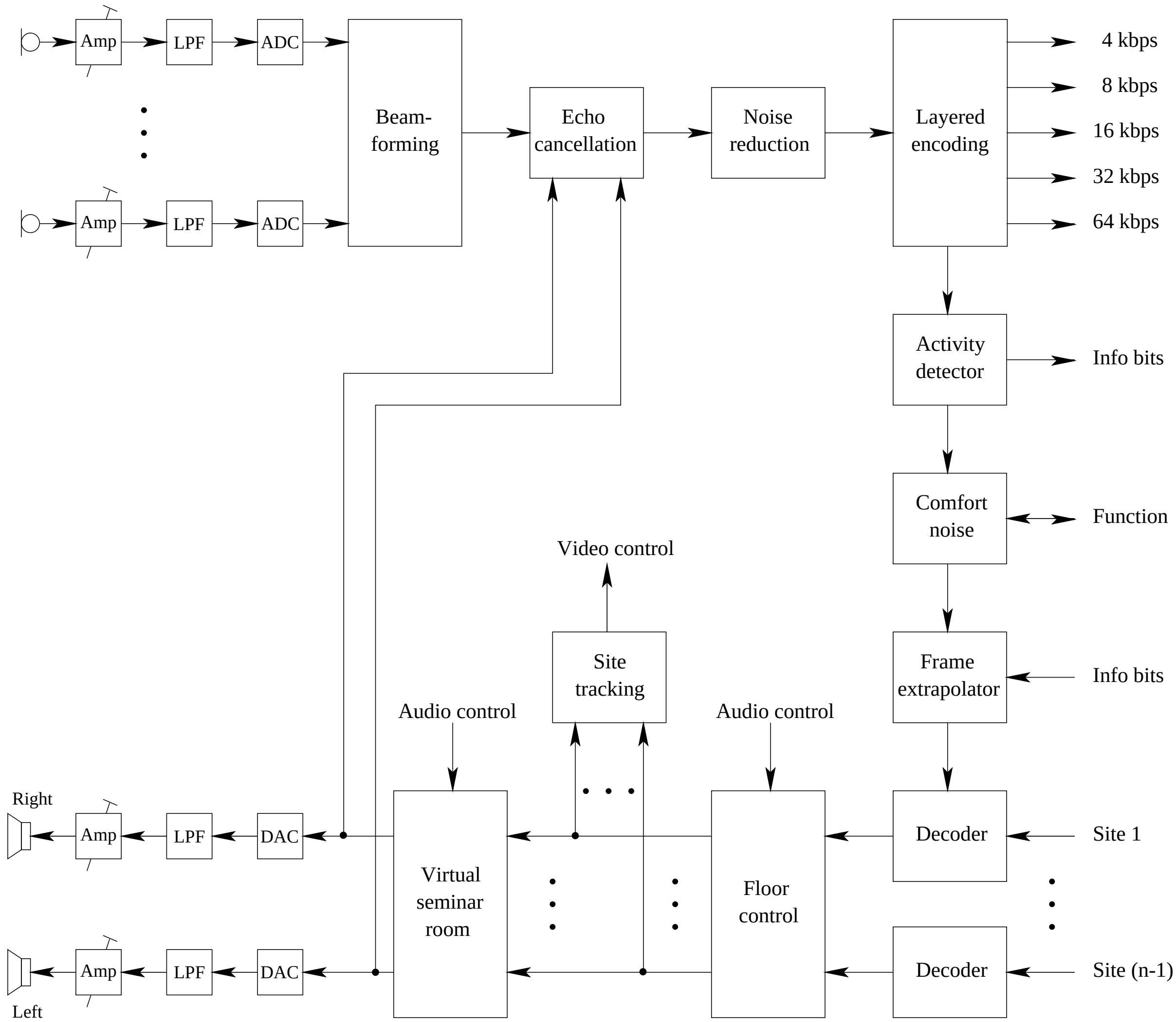


Figure 2.1 Layout of the audio part of the RTMM video conferencing system.

2.4 Interface

The architecture of the RTMM video conferencing system is shown in Figure 2.2, where the audio part, or audio agent, interfaces to:

- The network agent in form of encoded audio streams and info bits.
- The session control in form of synchronization, agent control, etc.

2.4.1 Synchronization

An important issue in the interface specification is the synchronization requirements and how to implement it. The discussion here is based on the survey by Blakowski and Steinmetz in [1].

Synchronization in multimedia systems refers to the temporal relations between independent media objects. The objects can be time-dependent like audio and video streams, where there

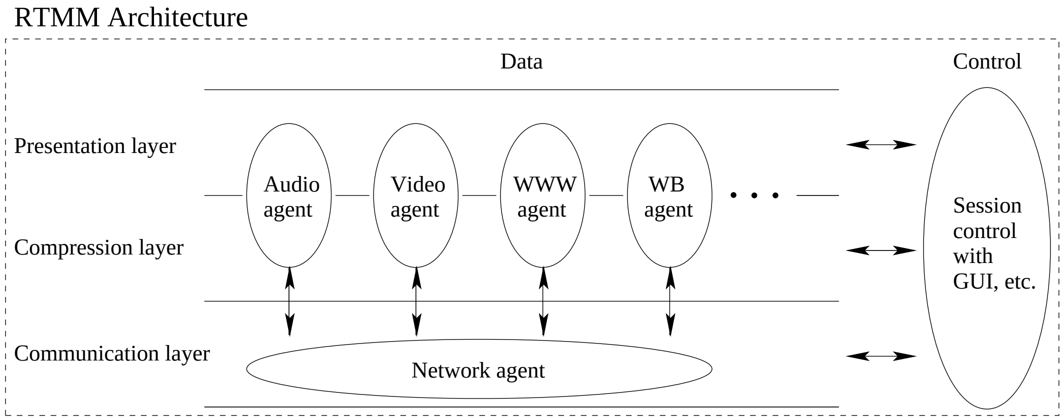


Figure 2.2 Architecture of the RTMM video conferencing system.

are temporal relations between consecutive units of the stream, or the objects can be time-independent like text and images on a wide-board.

The synchronization must be considered at several levels in a multimedia system. The operating system and lower communication layers handle single media streams with the objective to avoid jitter during the presentation, i.e., they must support *intra-object* synchronization (real-time operation). On top of this level is located the run-time support for the *inter-object* synchronization of multiple media streams, for example, the movements of the lips of a speaker and the related audio signal. Thus, the temporal relations between the media objects must be specified either implicitly during capturing of the media object, or explicitly in the case of presentations that are composed of independently objects.

2.4.1.1 Intra-object Synchronization

The audio stream is a time-dependent media object, which consists of a sequence of logical data units (samples), or since the sample duration is very small, a sequence of frames each containing a number of samples. In a live synchronization like the video conferencing application, the audio stream is created at the source, transmitted through the data path, and presented at the receiving site, i.e., the synchronization is specified implicitly during capturing. A possible manipulation is to adapt the presentation to the available resources, where it is preferable that such adaptions are performed at the source, especially in a distributed setup. Another type of live synchronization is to decouple the capturing and presentation of the media by an intermediate storage buffer, which allows the presentation speed to be changed.

Presentation requirements comprise, for intra-object synchronization, the accuracy concerning delays in the presentation or missing frames, which in both cases results in a gap in the stream, also known as the blocking problem. For audio streams, a simple solution is to stop the presentation and wait for the next frame, however, more acceptable solutions are to play the previous frame again, predict the next frame, or to resample the stream. The basic idea of resampling is to speed up or slow down streams for the purpose of synchronization. For audio signals, duplicated or deleted blocks, and changes in the playback rate can easily be noticed by the user, so more complex algorithms that can stretch or widen an audio sequence must be used.

One way to eliminate the blocking problem is to use QoS systems and/or intermediate storage buffers. In the RTMM system, the audio hardware and the communication layer will support QoS while the operating system will not as indicated in Figure 2.3. Thus, the operating system must specify with high probability the latency of service plus service time for moving a chunk of frames between the audio board and the network hardware. Together with the network delay, this service specification will give a lower limit of the presentation delay. For a given frame size, the service specification will also give a lower limit of the frame buffer size.

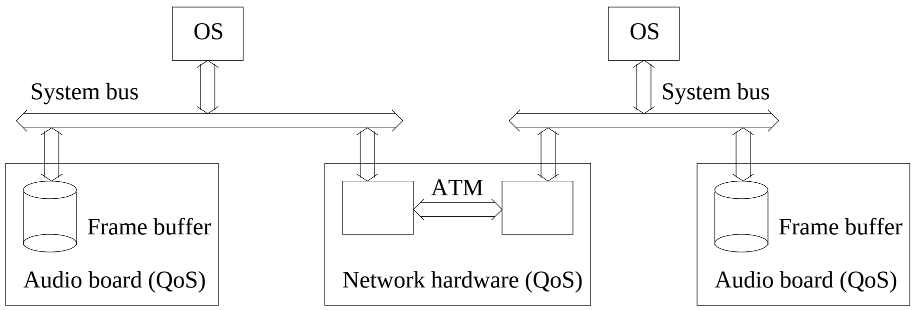


Figure 2.3 Synchronization with intermediate storage buffers.

2.4.1.2 Inter-object Synchronization

In video conferencing, the temporal relationship between the audio and video stream is refered to as lip synchronization, and the time difference is known as the skew. For inter-object synchronization, presentation requirements may be obtained from extensive experiments that are related to human perception. Such experiments shows that for a skew between -80 ms and +80 ms, most test candidates will not detect the synchronization error. Another observation is that video ahead of audio can be tolerated better than the opposing case.

2.4.2 Audio Protocol

2.5 Hardware

In order to obtain high quality sound, the audio agent should be implemented in dedicated hardware for professional audio applications. This includes microphones, mic pre-amps, sampling and processing board, power amplifier and loudspeakers. However, standard PC/workstation hardware should also be supported, i.e.,

- 1 microphone.
- Soundcard based I/O.
- Software based pre/post-processing and coding implemented such that the required CPU power and the resulting delay can be accepted.
- Two loudspeakers (stereo output).

In the following sections, the dedicated hardware used in this project are described.

2.5.1 Microphones

The microphones were chosen to have cardiod pick-up pattern based on the desired wall or ceiling placement, i.e., there will be effective rejection of the reverberations from the rear. Other critical microphone features are the sensitivity and noise level due to the long distance to the speaker.

Figure 2.4 Sennheiser K6 powering module with ME64 back-electret microphone capsule.

Thus, the microphones must be of the capacitor type, where the most cost effective solution was found to be the Sennheiser's K6 modular system[1]. Among the considered microphones were the AKG C480B/CK61-ULS[2], the Audio-Technica 4051a[3], the Neumann KM184[4], and the Shure BG4.1[5].

The K6 system consists of a powering module, which can be combined with various microphone modules to produce microphones with different pick-up patterns. The K6P powering module is for phantom power only, and features:

- Switchable bass roll-off filter protects against rumble, handling, pop and wind noise.
- On/off switch.
- High output, low noise.

and the chosen microphone module ME 64 features:

- Cardioid pick-up pattern.
- Excellent directivity across whole frequency range.
- Very good feedback rejection.
- Effective rejection of incidental rear noise.
- High speech clarity.
- High maximum sound pressure level.
- Wide frequency range.
- Integrated pop filter.

The microphone is shown in Figure 2.4, and the technical data for the modules can be seen in Table 2.1 and 2.2, respectively.

The microphones are placed in a custom-made array arrangement, where the dimensions can be chosen for maximum effectiveness and adaptability. The microphones are mounted into clamps and placed within two parallel tracks as shown in Figure 2.5(a). This arrangement allows for free motion of the microphones along the horizontal line. The distance between the two tracks can also be adjusted, and the microphones are acoustically decoupled from the tracks

[1] http://www.sennheiser.com
[2] http://www.akg-acoustics.com
[3] http://www.audio-technica.com
[4] http://www.neumann.com
[5] http://www.shure.com

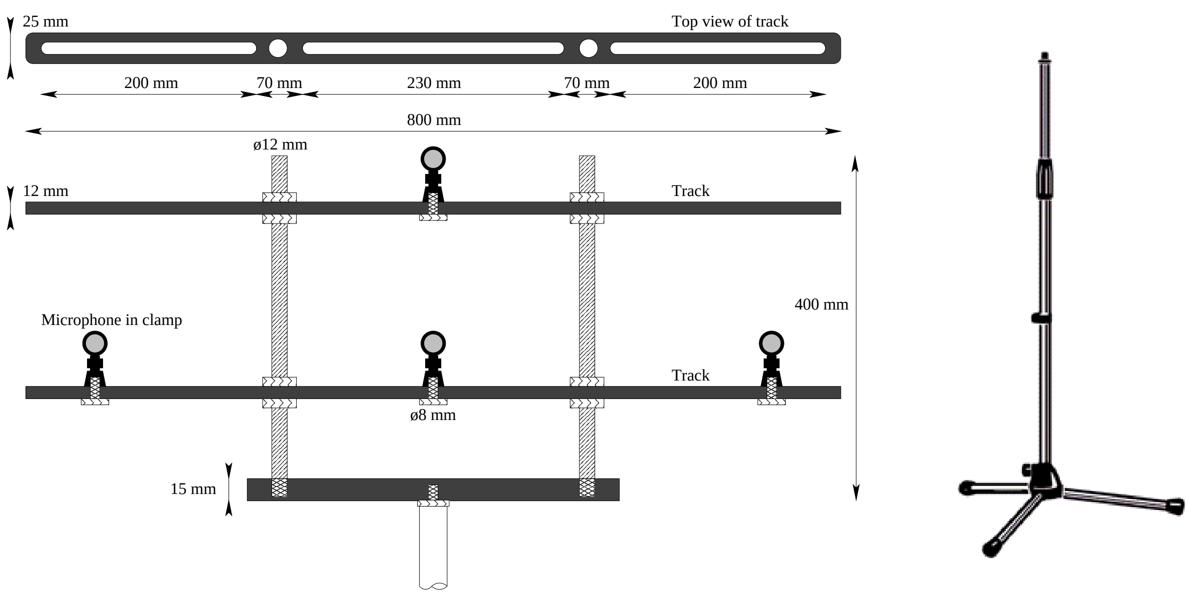
Frequency Respone	30 – 20,000 Hz ± 1 dB
Transmission Factor (1kHz)	0 dB
Nominal Impedance	200 ohms
Min. Terminating Impedance	1000 ohms
Max. Output Voltage	2 V (at 1 kHz)
Power Supply	Phantom 12 – 48 V
Supply Current	2 mA
Dimensions	Net: Ø 22 x 52 mm
Weight	30 g
Output Connector	Balanced XLRM-type

Table 2.1 Technical data for the Sennheiser K6P powering module.

Element	Back-electret
Pick-up Pattern	Cardioid
Frequency Respone	40 – 20,000 Hz ± 2.5 dB
Sensitivity (free field, no load, 1 kHz)	32 mV/Pa ± 2.5 dB
Equivalent Noise Level A-weighted	16 dB
Max. SPL (1 kHz)	130 dB (THD = 1 %)
Dimensions	Ø 22.5 x 106 mm
Weight	35 g

Table 2.2 Technical data for the Sennheiser ME64 back-electret microphone capsule.

by using vibration isolated clamps. The array arrangement can be mounted on a telescopic height-adjustable (89 to 160 cm) microphone stand[6] (see Figure 2.5(b)), or it can hang from the ceiling.



(a) (b)

Figure 2.5 (a) Diagram of the array system used for experiments. (b) Microphone stand.

[6]Microphone stand 201A/2 from Kønig & Meyer (http://www.k-m.dk).

Figure 2.6 TL Audio Ivory 5001 four-channel microphone pre-amp.

2.5.2 Microphone Pre-Amp

Microphone pre-amplifiers for professional standards are necessary to bring out the best in any quality microphone. The TL Audio Ivory 5001 model[7] was chosen as a cost efficient solution, and offers four top quality valve mic pre-amps in a single 2U package (see Figure 2.6). Each channel employs an ECC83/12AX7A twin stage triode valve, run from a stabilised 150V DC supply, and has input and output level controls, switchable +48V phantom power, phase reverse, 90 Hz low cut filter and "drive" and "peak" LEDs. Microphone inputs are on balaced XLRs, and Line outputs are provided on balanced XLR and unbalanced jack connectors.

The separate input and output level controls allow the valve stages to be driven harder if necessary for increased harmonic content, while still allowing the overall output level to be regulated. The drive LED monitors the input to the triode circuit to indicate the degree of valve "warming" that is being introduced, and the peak LED informing the user of impending clipping. Generally, it is advisable to set the input gain such that the peak LED just begins to illuminate on the loudest expected signal. This ensures that the optimum signal to noise ration is obtained, while allowing a good margin of headroom. The output fader permits anything from complete attenuation through to +15 dB of extra output gain – making it ideal for driving todays high level digital equipment. The technical data for the pre-amp can be seen in Table 2.3. Note, that most pre-amps in this class are based on valve technology, where another considered solution was the Behringer MIC2200[8].

2.5.3 Sampling and Processing Board

A multi-DSP board (PCI bus connection) must be available for real time pre/post-processing, encoding and decoding. The board must support multi-channel analog I/O, designed for audio applications, where at least 4 input and 2 output channels are available.

The chosen solution is the Spinner PCI format board from Bittware Research Systems[9], a high-performance, low-cost audio board designed for professional-audio OEM applications. The Spinner integrates 24-bit, 96 kHz analog and digital audio interfaces with Analog Devices' new low-cost ADSP-21065L SHARC chip. Thus, the Spinner goes well beyond CD-quality audio requirements, both in terms of sampling and resolution. The board is shown in Figure 2.7, the block diagram in Figure 2.8, and the specifications for the board are given in Table 2.4. The board features:

[7] http://www.tlaudio.co.uk
[8] http://www.behringer.de
[9] http://www.bittware.com

Microphone Inputs	Balanced
	Input impedance greater than 2 Kohm
	Gain range +16 dB to +60 dB
	Noise -127 dBu (EIN with 150 ohm source, 22 Hz – 22 KHz and maximum gain)
	3 pin female XLR connector
Phantom Power	48 V at 10 mA maximum per microphone
High Pass Filter	-3 dB at 90 Hz, 12 dB per octave
Balanced Outputs	Electronically balanced, unbalanced compatible
	Output impedance 47 ohms
	Maximum level +26 dBu, nominal level +4 dBu (1.2 Vrms)
	3 pin male XLR connector
Unbalanced Outputs	Output impedance 47 ohms
	Maximum level +18 dBu into 10 Kohms, nominal level -10 dBu (225 mVrms)
	2 pole 0.25" jack socket
Frequency Response	10 Hz to 40 KHz, +0, -1 dB
Dimensions	19" rack mounting, 2U high
	483 mm wide x 88 mm high x 200 mm deep
Weight	2.5 Kg

Table 2.3 Technical data for the TL Audio Ivory 5001 four-channel microphone pre-amp.

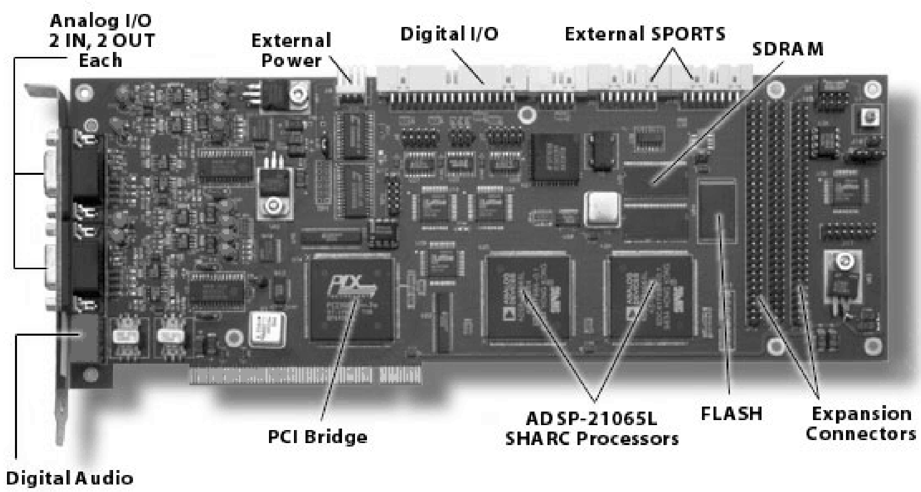


Figure 2.7 Spinner ADSP-21065L Audio OEM Board from Bittware.

ADSP-21065L Processor The Spinner is configured with one or two 60 MIPS ADSP-21065L processors, each with two bi-directional timers. Each processor also has twelve flags, eight of which are available for digital I/O via a digital I/O connector. The 21065L processors share a common 32-bit processors bus, which gives them access to the Spinner's 16MB bank of SDRAM (0-waitstate), 1M bank of FLASH memory, dual UART, and PCI bus interface.

Audio Interfaces The digital audio interface on the Spinner board consists of a single differential AES/EBU input and output channel. A digital audio transmitter and receiver transport AES/EBU signals to and from the 21065L processors via I^2S format serial ports. AES/EBU is the standard interface protocol that is used to transfer digital audio data between professional digital audio equipment such as PCM and DAT mastering recorders, modular multi-track recorders and other equipment.

Using the latest high-quality audio converters from AKM semiconductor, the Spinner's analog audio interface consists of two or four channels of A/D and two or four channels of D/A. The A/D and a D/A converters transmit the analog audio signals to and from the 21065L processors via I^2S format serial ports. The input channels are differential to enhance noise immunity.

Simultaneous sampling on all 4 analog inputs are very likely, since the A/D's are "paired" meaning there are 2-ch within a single IC, and both IC's in the 4-ch configuration receive the same master clock and word clock, so they are at least synchronous.

Spinner Options The Spinner is available with either one SHARC processor and an 8-bit digital I/O port for data acquisition, or with two SHARC processors and a 16-bit digital I/O port. The Spinner's 21065L processors can boot from the host via the PCI interface or operate standalone with the Spinner's on-board boot FLASH. The 1 MB FLASH is also available as non-volatile memory space. A 4M $\times$ 32 (16 MB) bank of SDRAM is available to the 21065L processors at a 60 MHz clock rate, and a dual UART allows the Spinner to communicate with external serial devices. Table 2.5 shows the configuration information for the Spinner board, where the board chosen here is the SPIN-6065-3.

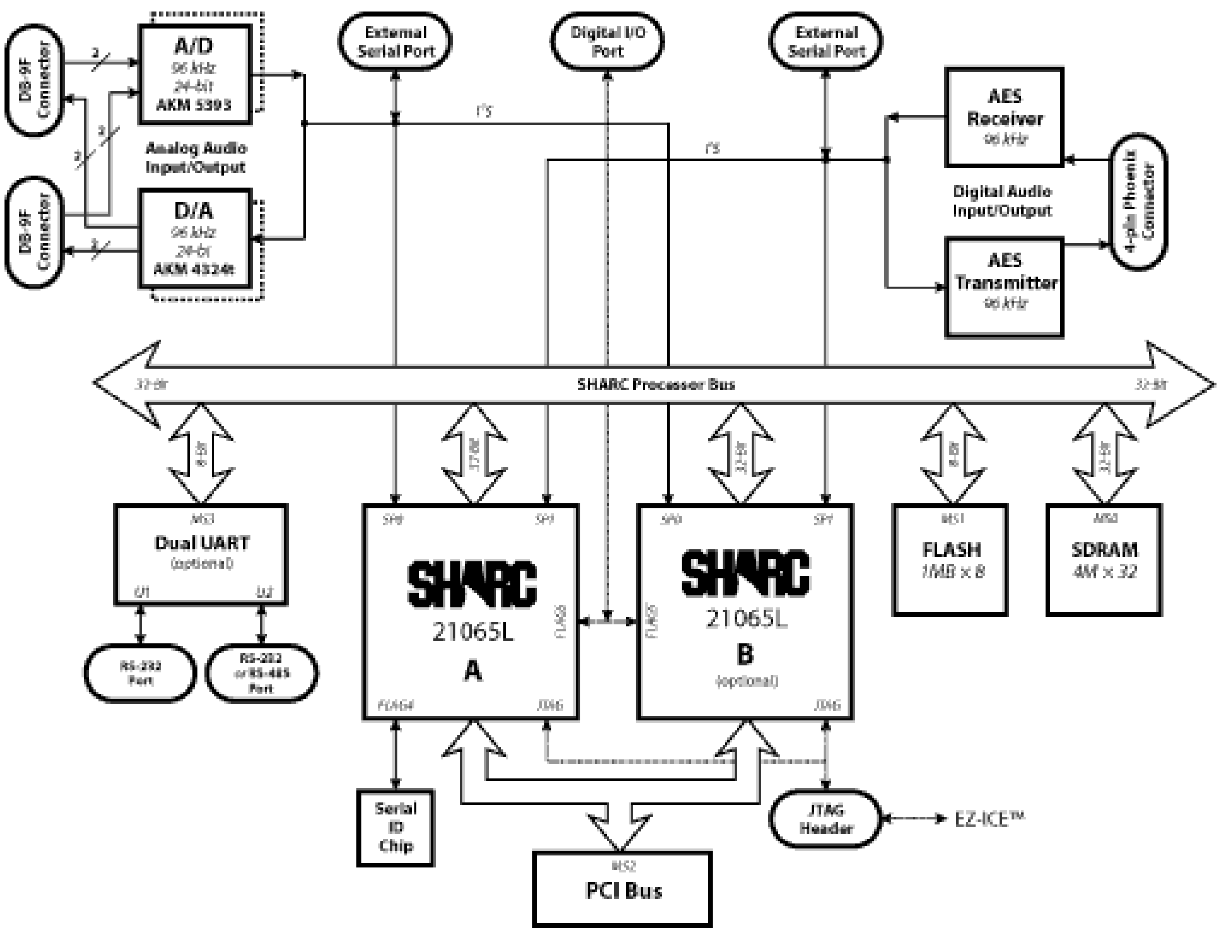


Figure 2.8 Spinner Block Diagram.

System	Processors	One or two Analog Devices ADSP-21065L SHARC DSPs
	External Memory	4M (16MB) × 32 SDRAM available to the 21065L at 60 MHz
		1M × 8 bank of FLASH memory for hostless boot or non-volatile data storage
	Serial Ports	Two external serial ports; one is connected to each processor
		One I^2S serial port per processor is routed to AES/EBU transmit and receive serial ports
		One I^2S serial port per processor is routed to the A/D and D/A converters
	Analog Audio Interface	24-bit 96 kHz A/D converter (AK5393) with two or four differential input channels
		24-bit 96 kHz D/A converter (AK4324) with two or four single-ended output channels
		Two DB-9 connectors, each with two input channels and two output channels
	Digital Audio Interface	24-bit 96 kHz audio transmitter supports one channel of AES/EBU data
		24-bit 96 kHz audio receiver supports one channel of AES/EBU data
		Supports I^2S serial port connection to the 21065L processor
		Four-pin Phoenix connector for AES/EBU I/O
	Digital I/O Interface	34-pin IDC connector provides 8 bits of digital I/O to each processor
	PCI Interface	Provides 32-bit master/slave (120 MB/s peak transfer rate) access to the SHARC processors
		All 21065L IOP registers are mapped to PCI memory space
		Supports hardware interrupts in both directions and host-based booting of the 21065L processors
	Dual UART	Two RS-232 ports (or one RS-232 port and one RS-485 port) are connected to a 16550 dual UART
	Debug Port	14-pin IDC header for IEEE JTAG 1149.1 boundary scan with extensions for in-circuit emulation
		Supports Analog Devices' EZ-ICE emulator
	Power	+5V @ 2A, +12V @ 200 mA, -12V @ 100 mA
	Size	10" × 4.2" PCI bus format
ADSP-21065L	Processing Rate	30 MHz, 180 MFLOPS peak, 60 MIPS
	Arithmetic	32/40-bit floating point, 32-bit fixed point arithmetic
	On-chip Memory	544 Kbits dual-ported on-chip SRAM
	Off-chip Addressing	64M words external address range
		SDRAM controller for glueless interface to external memory at 60 MHz
	I/O	Integrated I/O processor with four-channel DMA controller, twelve programmable I/O pins, two bi-directional timers, and two 30 Mb/s serial ports with I^2S support
Software	Host Interface	BittWare's Windows GUI-based software development kit for Windows 95 and Windows NT contains a C-callable library of board control and communications routines
		A porting kit is available for other operating system platforms
	Development Tools	Analog Devices VisualDSP tools include C compiler, assembler, linker, simulator, and debugger

Table 2.4 Specifications for the Spinner board.

Feature	SPIN-6065-1	SPIN-6065-2	SPIN-6065-3
Processors	1	1	2
Analog Audio I/O	2 ch @ 48 kHz	2 ch @ 96 kHz	4 ch @ 96 kHz
Digital Audio I/O	None	1 ch @ 96 kHz	1 ch @ 96 kHz
SDRAM	None	16 MB	16 MB
FLASH	None	1 MB	1 MB
UART	None	Dual RS-232	Dual RS-232
Digital I/O	8 bits	8 bits	16 bits
Standalone	No	Yes	Yes

Table 2.5 Configuration information for the Spinner board.

Software-Development Support BittWare's complete Windows GUI-based software development kit allows users to easily integrate the audio board into their systems. Available software tools include a new integrated development environment with a built-in debugger, BittWare's updated DSP21k Toolkit with a comprehensive host interface library and driver support, and Analog Devices' VisualDSP software development tools. The Spinner also supports in-circuit emulation with Analog Devices' EZ-ICE emulator and is enhanced by a variety of third-party support applications.

Boards in Parallel The sampling can be synchronized between two or more boards in parallel by using the AES/EBU clock as the A/D sample clock. So a "reference" clock can be distributed using the AES/EBU input. Communication between several boards can be done in the following ways:

- The UART is one way.
- The SHARC serial ports can be used, but not if no A/D-D/A one external SPORT is available, and if no digital audio the other SPORT is available.
- PCI bus mastering.
- Auxillary FLAGs and IRQs.
- The entire system bus will be added to expansion connectors (compatible with the EzKit) to the production release.

Finally, it should be mentioned that the chosen board is the only low-cost solution at the moment. The products at the marked are typically general purpose multi-DSP boards like the BlackTip-MCM ISA format board[10] from Bittware Research Systems, based on the Analog Devices AD14060 Quad-SHARC Multichip Module. Multi-channel I/O is then obtained by a daughter board designed for audio applications, e.g., the 8-channel Bitsi Audio I/O Mezzanine site also from Bittware Research Systems. The problem is that this combination cost like 8 times the Spinner board. Another obvious solution is to use multi-channel recording/sampling products made for the home-studio marked, e.g., the Layla from Echo[11]. However, these cards are based on standard high-level software, and do not support SDK for low-level operation.

[10] http://www.bittware.com
[11] http://www.echospeech.com

Power Amplifier	Continuous average power into 8 ohms	50 W (17 dBW)
	Rated distortion (THD 20 Hz – 20 kHz)	0.03 %
	Clipping power (max. continuous power/channel)	55 W
	IHF dynamic headroom at 8 ohms	3.5 dB
	IHF dynamic power - 8 ohms	100 W (20 dBW)
	- 4 ohms	130 W (21.1 dBW)
	- 2 ohms	170 W (22.3 dBW)
	Damping factor (ref. 8 ohms 50 Hz)	> 60
	Signal/noise ratio, A weighted - ref. 1 W	90 dB
	- ref. rated power	104 dB
Line Level Inputs	Input impedance (R and C)	20k ohms / 500 pF
	Input sensitivity ref. rated output	165 mV
	Frequency response, 20 Hz – 20 kHz	±0.3 dB
Line Level Outputs	Output impedance - Pre-amp	220 ohms
	- Tape	Source Z + 2k ohms
	- Phones	220 ohms
Controls	Treble	5 dB at 10 kHz
	Bass	7 dB at 100 Hz
	Remote control	Yes
Physical Spec.	Dimensions in mm (WxHxD)	435x110x290
	Net weight	7.6 kg

Table 2.6 Technical data for the NAD model C340 integrated amplifier.

2.5.4 Power Amplifier

The power amplifier is not so critical, and a quality, mid-price amplifier like the NAD model C340[12] can do the job (see Figure 2.9). The C340 has 7 line inputs (including 2 tape in/outputs with dubbing facility) and the pre-amplifier section can be separated from the power amplifier for easy upgrades (bi-amping, for instance) or adding ancillary equipment without making the C340 redundant. The headphone socket will drive virtually any non-electrostatic headphone. The C340 tone controls only work at the frequency extremes leaving the critical mid-band essentially unaltered. The tone control circuits can be completely bypassed by using the tone defeat switch. The technical data for the amplifier can be seen in Table 2.6.

Figure 2.9 NAD model C340 integrated amplifier.

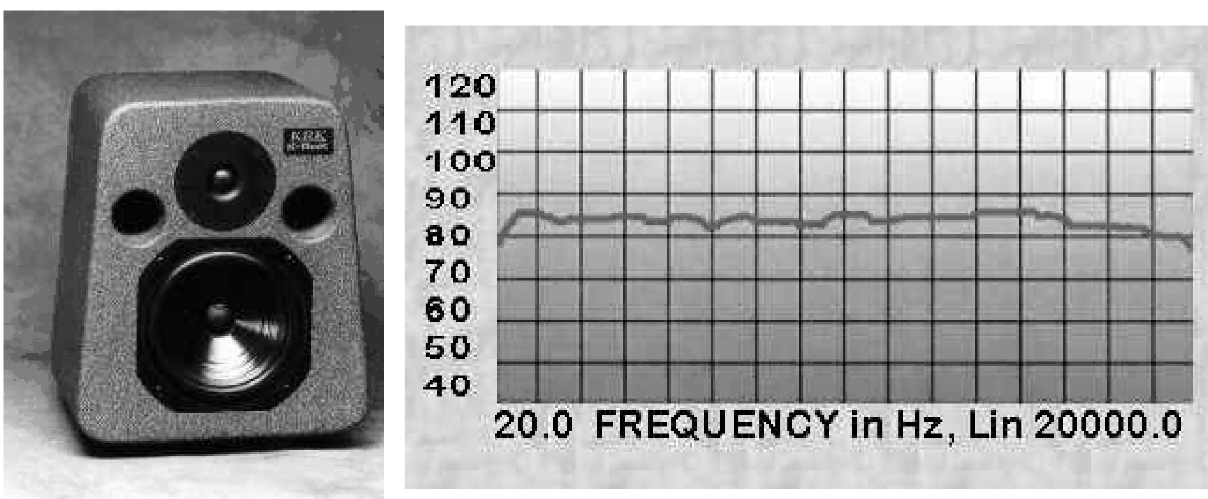
[12]http://www.nad.co.uk

Drivers	7" latex-coated woofer; 1" silk-dome tweeter
Crossover	Passive 2-way
Crossover Point	2.5 kHz
Frequency Response	± 3 dB, 57 Hz – 19 kHz
Sensitivity	92 dB (1 watt @ 1 meter)
Max. Power Handling	100 watts
Max. SPL	106 dB (@ 1 meter)
Nominal Impedance	8 ohms
Cabinet Dimensions	H-14", W-12", D-9.75"
Cabinet Finish	Custom gray texture
Shipping Weight	36 lbs. per pair

Table 2.7 Technical data for the KRoK studie monitor.

2.5.5 Loudspeakers

The two loudspeakers must be of studio quality, and the cost effective KRoK 7" 2-way studio monitor from KRK Systems[13] was found to be a good choice (see Figure 2.10(a)). The loudspeaker features low distortion, high power handling, superb imaging, and extremely smooth frequency response (see Figure 2.10(b)). Magnetic driver shielding enabling the speakers to be used near computer monitors without compromising picture quality, however in this setup, the speakers are placed on stands. The technical data for the speakers can be seen in Table 2.7. Other considered loudspeakers were the Yamaha NS10M Studio[14], and the Alesis Monitor One[15].



(a) (b)

Figure 2.10 (a) KRoK studie monitor from KRK Systems. (b) Frequency response graph.

[13]http://www.krksys.com
[14]http://www.yamaha.co.uk
[15]http://www.alesis.com

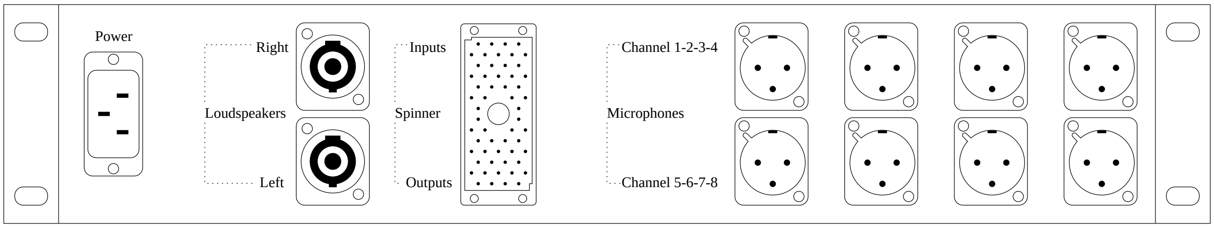


Figure 2.11 Front panel connections in rack-mount installation.

2.5.6 Rack-Mount Installation and Cables

To have a movable and easy to use system, the pre- and power amplifiers are mounted in a standard SKB 19-inch audio rack case[16], where all input/output connectors are hardwired to the front panel shown in Figure 2.11. The front panel contains eight microphone XLR-F input connectors, a 56-way EDAC receptacle with eight I/O connections to the Spinner board, and two Speakon connectors (female) to the loudspeakers. The above mentioned equipment are connected using cables/plugs as listed below:

- The microphone to preamp (XLR front panel) connections are Cordial Pro Line CPM 10 FM[17] cables, which is a 10 meter XLR-XLR black cable with Neutrik NC 3 FX/MX XLR plugs.
- The cable that connects the Spinner boards DB-9F dual analog inputs/outputs with the 56-way EDAC receptacle on the front panel is custom-made, i.e., two Spinner boards (four SUB-D connectors) are connected to the mic. pre-amps and the power amp by using a cabel with 16 shielded pairs. One I/O connection is shown in Figure 2.12.
- The loudspeaker caples are 10 meters of the 2-pole Cordial cable CLS 225-651, mounted with standard speaker bunch pin plugs in one end, and Speakon connector (male) in the other end. This 2.5 mm^2 loudspeaker-cable is robust, easy to reel and made for universal use. The extremely high flexibility is among others achieved by the very low diameter of the strands of 0.07 mm.

Note, that Neutrik[18] set the state of the art in XLR and Speakon audio connector technology. The pinout of balanced XLR, EDAC, and SUB-D connections are illustrated in Figure 2.13.

2.5.7 Supplier and Price Information

Information about the danish suppliers and prices for the considered system components are shown in Table 2.8. The total price is for four microphone input, however, the price for adding another four microphone block to the system is also given. The prices for the software packages are listet in Table 2.9.

2.6 Setup

The physical setup of the RTMM video conferencing system could be the one shown in Figure 2.14, where the microphone array could be placed either above the (main) camera as indi-

[16] http://www.skbcases.com
[17] http://www.cordial-gmbh.de
[18] http://www.neutrik.com

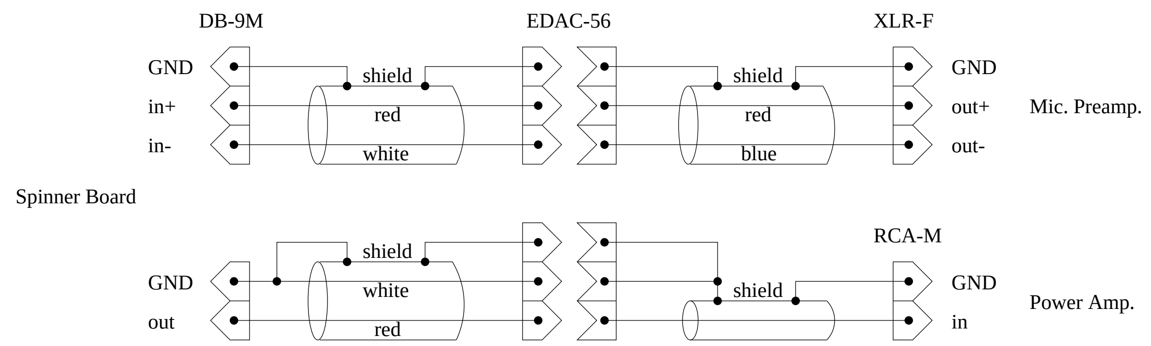


Figure 2.12 One I/O cabel connection between Spinner board and amplifiers.

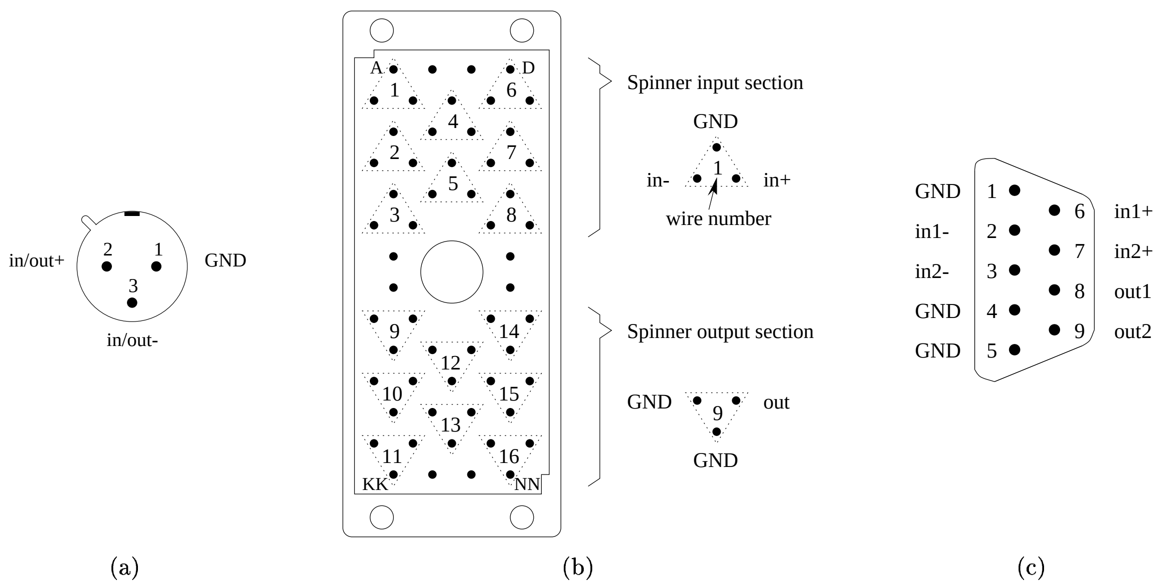


Figure 2.13 (a) Rear view of XLR-M or front view of XLR-F connector. (b) Rear view of EDAC-56 plug or front view of EDAC-56 receptacle. (c) Rear view of DB-9M or front view of DB-9F connector (Spinner board).

cated in the Figure or in the center of the room. The latter could be on the conference table or hanging from the ceiling. Thus, both desktop environments and more specialized seminar rooms are supported.

Product	Supplier	Units	Unit price	Total price
Sennheiser K6P mic power module	Kinovox	4	1289	5156
Sennheiser ME64 mic capsule	Kinovox	4	1184	4736
Kønig & Meyer 201A/2 mic stand	Kinovox	1	288	288
Array arrangement	Custom-made DTU	1	1850	1850
Cordial Pro Line CPM 10 FM mic caple	Kinovox	4	169	676
TL Audio Ivory 5001 mic pre-amp	Musikhuset	1	4230	4230
Four DB-9M to EDAC-56 cable (5 meter)	Custom-made	1	580	580
Bittware Spinner SPIN-6065-3 board	Metrix	1	11457	11457
NAD model C340 integrated amplifier	Hi-Fi Klubben	1	1999	1999
10 meter Cordial CLS 225-651 loudspeaker cable	Kinovox	2	161	322
KRoK studio monitors from KRK Systems	Kinovox	2	2338	4676
Loudspeaker stands	Ikea	2	72	144
Rack with hardwired front panel	Custom-made	1	3050	3050
System price (4 mic)				39164
Additional price for blocks of 4 mic				26255

Table 2.8 Supplier and price information (Danish kroner without tax, February 1999. Note, that the prices from Kinovox are after 10 % reduction given to institutions).

Product	Supplier	List price
Spinner board support package	Metrix	2710
VisualDSP software development tools (version 4.0.1)	Analog Devices	18840

Table 2.9 Software price information (Danish kroner without tax, February 1999).

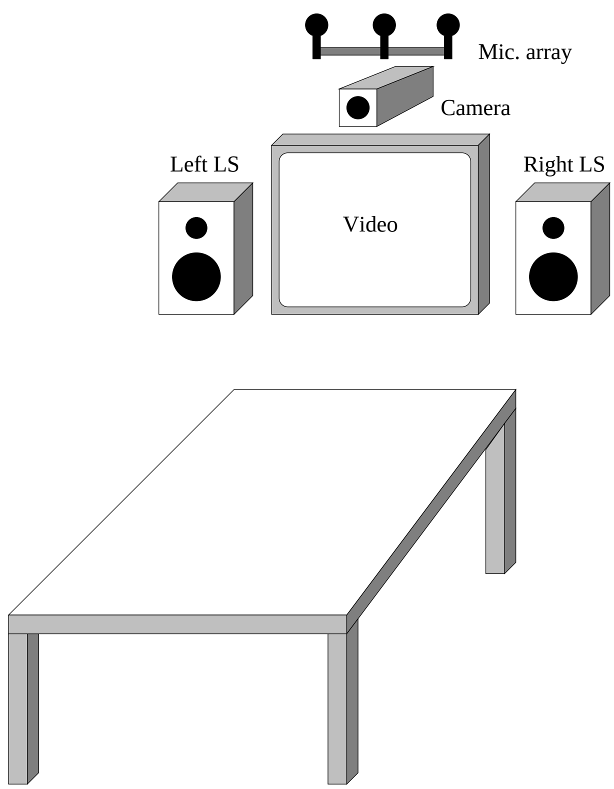


Figure 2.14 Setup of the RTMM video conferencing system.