AXIS 2191 Audio Module

Sight, Speech & Sound — over IP

Last update: Nov 13, 2001

TABLE OF CONTENTS

1 Introduction

The AXIS 2191 Audio Module is a hardware accessory that provides audio capability for Axis network cameras. This add-on module connects the serial port of the host network camera. The product combination delivers live audio with quality moving images that are experienced directly from an Internet Explorer Web browser.

This paper describes the AXIS 2191 Audio module from a technical perspective and explains how the product works. It also describes suitable application areas for the product and briefly discusses future development possibilities using the API.

For additional information, please refer to the official product datasheet available on our website at: http://www.axis.com

2 Why an audio module?

The importance of audio has increased in parity with the increasing popularity of independent network cameras. With many end users demanding audio as a supplement for their video applications, the introduction of two-way sound for Axis video products will add even greater dimension to the huge variety of existing remote monitoring and Web attraction applications. With the AXIS 2191 Audio Module, users can now monitor and communicate with remote and local sites using Internet Explorer from their workstations.

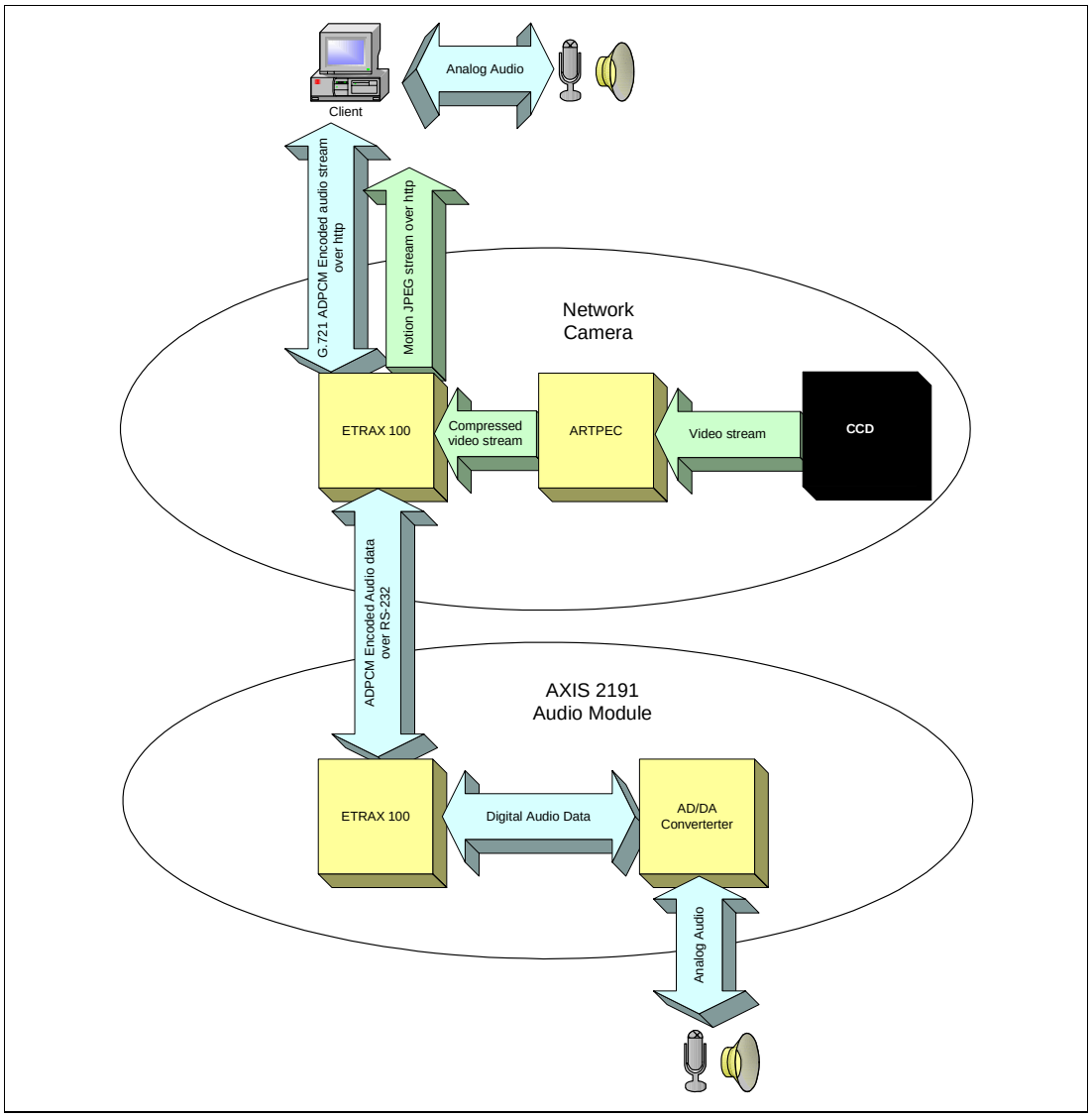
Through the ADP Program (Axis Developer Partner), Axis continues with a commitment towards helping third-party developers integrate this new audio-video functionality into the new generation of audio-enabled network applications.

3 Theory of Operation

The AXIS 2191 Audio Module is designed for use strictly in conjunction with Axis network cameras and cannot be used as standalone units. It connects to the Axis network camera via the RS-232 serial port using the supplied null-modem cable. Delivering a G.721 encoded audio stream, the connecting camera is then able to redistribute the audio signal to one or several recipient(s) using its built-in Web server and the HTTP protocol.

Clients will automatically receive the required Active-X component through their Internet Explorer Web browser. The installation of this component onto the client workstation facilitates the reception of video as well as sending and reception of the audio stream.

3.1 Audio & Video Data Flow



The flow of audio through the audio module and connecting network camera (illustrated above) can be described as follows:

1. The analogue audio signal is received from either the built-in or external microphone. This signal is then converted by the AD/DA (Analog Digital/Digital Analog) converter into uncompressed a 16-bit digital audio-data video stream with a sampling rate of 8 KHz.
2. The AXIS ETRAX CPU within the Audio module compresses the received audio using a software coder/decoder (codec) to G.721 ADPCM (See Glossary of terms) format and sends the data to the serial port of the connecting network camera using RS-232.
3. The network camera sends encoded data to the client in a separate stream using HTTP networks. This data is in transmitted in parallel with the motion JPEG image data received from the ARTPEC-1 chip.

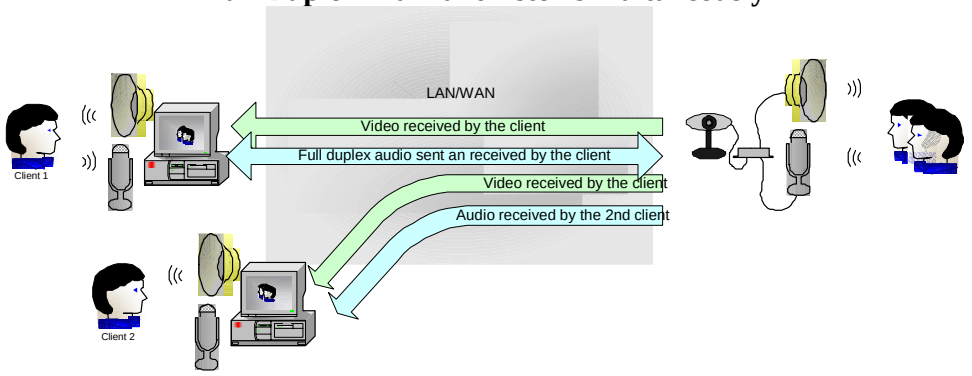
4. The client workstation receives the audio stream and decompresses it using the installed Active-X component and a software codec. The audio signal is then distributed to the Windows multimedia sub-system to be heard through the workstation speakers via the installed soundcard of the PC.

Note that the audio can be sent from the client to the audio module through the camera in a reverse direction: The clients audio data is directed to the network camera .The network camera then redistributes the digital audio signal to the audio-module for decoding prior to being heard by the receiver via the AD/DA converter and loudspeakers.

3.2 Audio Modes

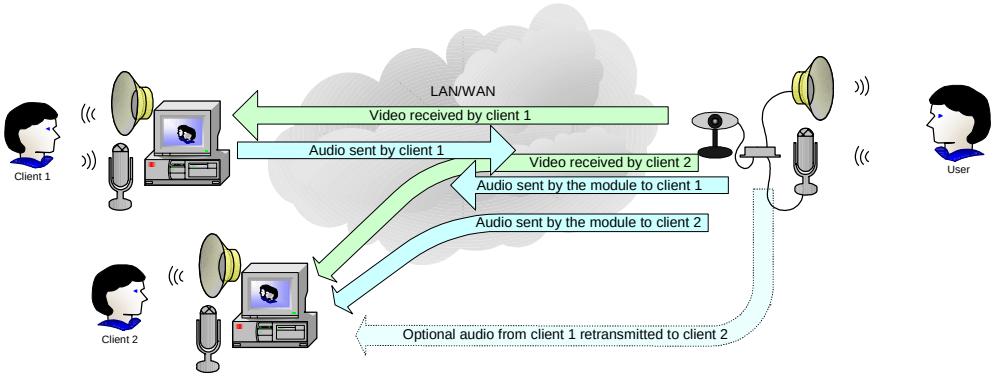
Following four audio modes are possible:

Full Duplex -Talk and listen simultaneously



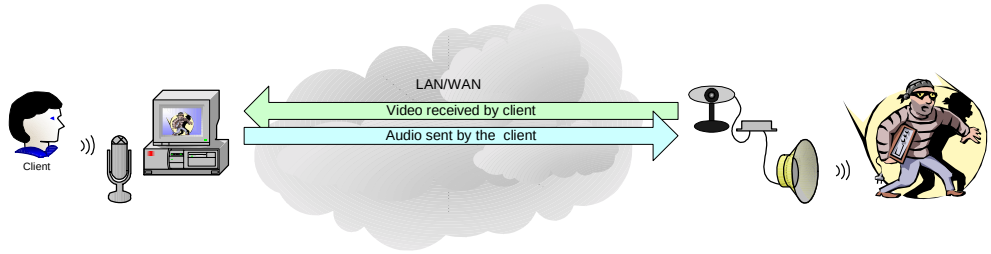
Applicable for remote meeting applications (note bandwidth requirements). Only the first connecting client will be able to send audio, and of course receive audio and video. Other connecting clients will be able only to receive audio and video.

Half Duplex - Talk and listen



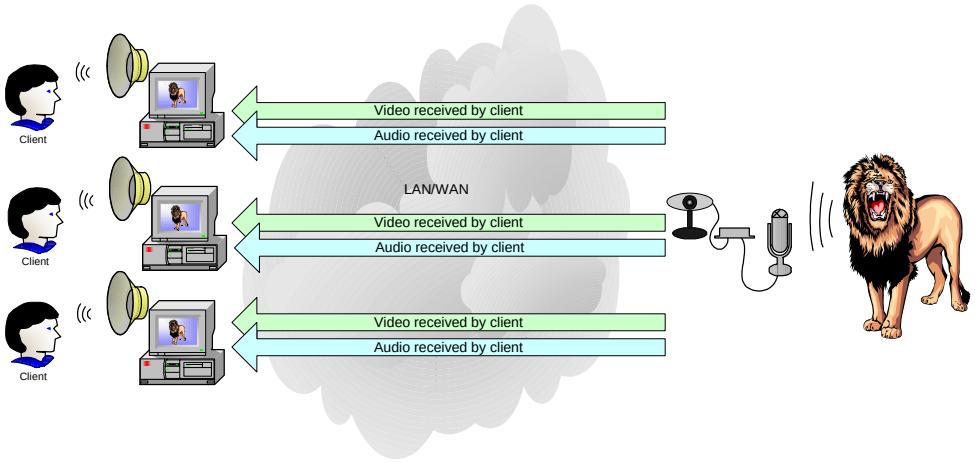
This mode is applicable for remote meeting and remote monitoring applications— such as entrance systems. In this mode, a special setting is available for re-distributing the audio stream from one client to another. This setting can prove particularly useful where single speaker control is desirable, i.e. within remote guidance scenarios.

Simplex – Talk to the module



Applicable for the remote monitoring, e.g. entrance camera applications.

Simplex – Listen to the module



Applicable for the remote monitoring and Web attraction applications, where the ability to hear is of the highest priority.

4 Applications

The audio module has been primarily developed towards the following target applications.

Many different organizations are now utilizing their corporate networks in new ways. In remote monitoring applications companies can now use Axis network cameras to keep a better visual check on their property, technical equipment, or any other assets that they value.

With travelling overheads compounding the professional costs for diagnosing and repairing defective technical equipment off-site, the ability to see and remotely communicate with specialists located onsite can provide tremendous savings opportunities against any organizations field service costs.

4.1 Remote Monitoring

- **Door camera/communication system**
  - Use the camera/audio module to view and communicate with the visitor before opening the door using the output function of the camera.
- **Audio monitoring**
  - View and listen to people entering sensitive areas within your company
  - Speak to and challenge unauthorized visitors attempting to enter sensitive areas

4.2 Web Attraction

- **Web attraction**
  - Listen to live audio in parallel with the live video. The added functionality enhances the information content.

The fascination of being able, not only to see, but to hear what is happening in any remote area of the world rarely fails to stimulate the curiosity of the Internet users. Streaming live video that captures the color and vibrancy of a city's nightlife or bustling marketplace is a virtual experience that the majority of Internet surfers will never grow tired of.

Axis network cameras, together with the AXIS 2191 Audio Module provides Website developers with a unique opportunity for creating live, audio-visual, Web pages over the Internet. Empowered and encouraged by the continued rollout of Axis video technology, Internet users have the potential to surf-in and enjoy a live audiovisual experience from any place of popular interest – be it a local town center, exotic tourist resort, or any public meeting place. Please note the bandwidth requirements for these applications.

4.3 Remote Meeting and Guidance

- **Remote Meeting**
  - Convenient and cost effective way of participating in remote meetings without being physically present. A remote participant can use a PC together with Internet Explorer and be "virtually present" saving on travel expenses.
- **Remote guidance/assistance**
  - A specialist can guide users and colleagues on remote site locations with the help of live audio/video - the specialist can be "virtually present" at any remote site.

Using an Axis network camera and an AXIS 2191 Audio module, business people need only type the IP address of a remote camera into the browser to communicate visually and audibly with anybody – in any remote office –that they wish!

With the AXIS 2191 there is now no real need to schedule future business meetings. Now it is possible to surf-in for a casual meeting. Simply navigate the way to the

meeting in the same way that any Internet site is browsed- via the URL in the browser. With Axis network cameras and the new Axis audio module it is possible to enjoy instant video, and instant two-way audio, through Microsoft Internet Explorer. Please note that this is the only browser compatible with the Audio Module.

Important Note! The bandwidth limitations on your local network will restrict the number simultaneous audio-visual client connections that can be maintained without degradation of the audio-visual quality. Axis strongly advises to check the local bandwidth restrictions on the network prior to setting up the application. The following section provides more detailed discussion concerning this important consideration.

5 Bandwidth Requirements

To understand the bandwidth required for successful usage of the AXIS 2191 Audio Module together with an Axis networks camera, one first has to consider why Motion JPEG is a preferred image format for transferring video streams over TCP/IP networks. Network and system managers that are developing remote monitoring applications will put great emphasis on sourcing video products that can consistently provide images of high quality. Unlike H-compression and MPEG image formats, the Motion JPEG format supported within Axis' video products provides optimal image quality for every individual image— within any given image sequence. Furthermore, this format is native to most standard Web browsers and can be easily imported into the vast majority of Web-based applications.

The superior quality of the pictures has close coupling to the size of the pictures, as illustrated below:

Resolution	**Average Compression level**		
	Low	**Medium**	**High**
704 x 576 (50 Hz)	52kb	34kb	20kb
352 x 288 (50 Hz)	12kb	8kb	4kb
704x480 (60 Hz)	36kb	22kb	13kb
352x240 (60 Hz)	8kb	4kb	3kb

5.1 Calculating Required Bandwidth

The bandwidth required for transferring audio/video data over the network can be calculated by using the load demanded by one client and a low compressed average picture (12 Kb) with the resolution of 352 x 288.

Frames/s	Required bandwidth video stream only (Mbit/s)	Required bandwidth video stream & half duplex audio (Mbit/s)	Required bandwidth video stream & full duplex audio (Mbit/s)
1	0.096	0.135	0.160
2	0.192	0.231	0.256
3	0.288	0.327	0.352
4	0.384	0.423	0.448
5	0.480	0.519	0.544
6	0.576	0.615	0.640
7	0.672	0.711	0.736
8	0.768	0.807	0.832
9	0.864	0.903	0.928
10	0.960	0.999	1.024
11	1.056	1.095	1.120
12	1.152	1.191	1.216

Note: Above table is reflecting the bandwidth requirements for one connected client. The required bandwidth is then increasing with the same value for each additional client.

The audio functionality adds a minimal overhead to the total data stream. The audio always requires a total of 39,2 kbit/sec for half duplex, for full duplex the value is 71,2 kbit/sec (2 x 32 kbit/s audio data + 7.2 kbit/s (header information))

6 API

Third party developers may take advantage of the Axis HTTP API to develop applications for the network camera/audio module combination, more information for developers can be found on
Axis' website at: http://www.axis.com/techsup/cam_servers/dev/index.htm

7 Technical Specifications

- Operating temperature: 40-105oF (5-40oC).
- Humidity - 8-80% relative humidity.
- EMC - CE EN55024, EN55022, Class B and EN61000-3-3.
- EMC - FCC Class A of FCC Rules and Regulations part 15, subpart B.
- EMC -
- Full-duplex audio: Audio data encoded
- ADPCM format at 32kbps, 8 kHz sampling (G.721). Data is sent using HTTP.
- 9-pin D-SUB serial connector: RS-232.
- Power Input: Axis PS-D power supply.
- Microphone Input: 1-50mVpp. PC type.
- Line Out: Unbalanced, 0.05-1.0Vpp.
- Line Input: Balanced 0.05-1Vpp. Connect source ground to pin 2 and source signal to pin 1 if the source is unbalanced.
- Speaker Output: Balanced, 0.5W. Impedance 8-32 Ohms. Connect directly to speaker without capacitors.
- Alternative Power: 12-15VAC, min 10VA, or 15-20VDC, min 7W.
- Metrics: Height: 1.1” (27mm), Width: 4.4” (112mm), Length: 4.3” (110mm),
- Weight: 0.7lb (0.32kg).
- Maximum number of users: 10 (on local area network).

8 Glossary of Terms

8.1 ADPCM

ADPCM, *Adaptive Differential Pulse Code Modulation,* is a standard method for converting analog signals to digital signals; where, the PCM signal is a quantified version of the actual analog signal. ADPCM only records the difference between two measurements and therefore uses less bandwidth. It takes advantage of the fact that you can generally predict the value of the next sound sample based on the previous sound sample.

8.2 ARTPEC-1

Axis proprietary image compression chip ARTPEC-1 is used to compress the video stream received from the *CCD* to JPEG.

8.3 CCD

Charge Coupled Device, image sensor.

8.4 CCITT

CCITT Consultative Committee for International Telegraph and Telephone, now known as the ITU-T (the Telecommunication Standardization Sector of the International Telecommunications Union).

8.5 Codec

Short for compressor/decompressor, ***Codecs*** are various types of algorithms that are applied to audio, video, and image data. The benefit of using these algorithms is that the resultant audio-video data streams demand much less network bandwidth. The same *codec* used for compressing data must be used for decompressing of the data.

8.6 ETRAX-100

ETRAX 100 is Axis' fifth-generation RISC CPU. It is an optimised system-on-a-chip solution for putting peripherals on the network. The ETRAX 100 was developed using ASIC technology. The innovative 100 MIPS 32-bit RISC design delivers compact code and exceptional performance at low power consumption. An 8-kbyte on-chip cache helps to take full advantage of the CPU performance.

8.7 G.721

G.721 is a CCITT defined public standards for compressing voice data in using ADPCM at 32 kbits/sec.

8.8 JPEG

Joint Photographic Experts Group, standard for image compression.