

Audio in Axis' video products

TABLE OF CONTENTS

- **1. OVERVIEW**
- **2. AUDIO CODECS**
- **3. AUDIO HTTP API**
 - **3.1 Audio MIME types**
 - **3.2 Audio data request**
 - **3.3 Singlepart audio data response**
 - **3.4 Multipart audio data response**
 - **3.5 Audio data transmit**
- **4. AUDIO PARAMETERS**
 - **4.1 Audio**
 - **4.2 Audio.A#**
 - **4.3 AudioSource**
 - **4.4 AudioSource.A0**
 - **4.5 Properties.Audio**
- **5. AXIS ACTIVEX COMPONENTS**
- **6. AUDIO SETTINGS**
 - **6.1 Audio modes**
 - **6.2 Max number of clients**
 - **6.3 Broadcast the audio from the active client**
 - **6.4 Audio input**
 - **6.5 Audio output**
 - **6.6 Audio quality settings**
- **7. CLIENT COMPUTER SETTINGS FOR SENDING AUDIO TO THE CAMERA**

1. OVERVIEW

Axis video products with built-in audio support let remote users see, hear and speak over IP networks. These products are suitable for surveillance and remote monitoring applications that require professional video and two-way audio communication over networks.

With a built-in microphone and the ability to connect to external speakers, these products enable remote users to not only view, but also listen in on an area and communicate orders or requests to visitors or intruders. These products are ideal choices for securing offices, shops and other facilities such as schools, university campuses and prisons.

Network cameras with audio support can be found here:

<http://www.axis.com/products/video/camera/productguide.htm>

Video servers with audio support can be found here:

http://www.axis.com/products/video/video_server/productguide.htm

Note however, that this document only covers products using firmware version 4.00 and above. A list of Axis' video products and accompanying firmware is found in the Product Interface Guide:

http://www.axis.com/techsup/cam_servers/dev/product_interface_guide.htm

IMPORTANT NOTICES!

Axis Communications AB provides **no** guarantee that any of the examples shown in this document will work for any particular application.

Axis Communications AB **cannot** and **will not** be held liable for any damage inflicted to any device as a result of the examples or instructions mentioned in this document.

Axis Communications AB reserves the right to make changes to this document without prior notice.

Please bear in mind that the flash chip manufacturer estimates the number of writes to the flash chips to about 100,000. Writing a lot of temporary files to the flash memory should thus be avoided. Use the ram disk mounted on /tmp instead.

2. AUDIO CODECS

A codec performs media-data compression and decompression. When a track is encoded, it is converted to a compressed format suitable for storage or transmission, when it is decoded it is converted to a non-compressed (raw) format suitable for presentation.

Axis video products with firmware version 4.xx support G.711 μ -law, G.726 24 and 32 kbit/s.

More information about the codecs and ability to FTP download code for the encoders/decoders from Sun Microsystems, Inc is available here:

http://www-mobile.ecs.soton.ac.uk/speech_codecs/standards/adpcm.html

3. AUDIO HTTP API

When Axis introduced the new generation of Axis video products, a new version of the HTTP API was introduced to cover all new functionalities in these products. The new generation video products, with firmware version 4.00 and above, are in most cases backward compatible and still support the Axis HTTP API version 1, but the HTTP API version 2 should be used to be able to use all the new features in these products.

The audio HTTP requests are almost the same in the HTTP API version 1 and version 2. There is however a new parameter available in HTTP API version 2, the httptype parameter, that can be used to choose the audio streaming method, multipart or singlepart audio. The param.cgi that was introduced in HTTP API version 2 to handle parameters can be used to get and set the audio parameters but the /axis-cgi/audio/getparam.cgi requests in the HTTP API version 1 will continue to be supported by these products for some time to come. The response differs since the audio parameters have been changed.

3.1 Audio MIME types

Supported MIME types for audio

audio/basic	which is G.711 μ -law 64kbit/s
audio/32KADPCM	which is G.726 32kbit/s
audio/G723	which is G.726 24kbit/s

3.2 Audio data request

Request an audio stream.

Method: GET

Syntax:

```
http://<servername>/axis-cgi/audio/receive.cgi[?<parameter>=<value>]
```

with the following parameters and values

<parameter>=<value>	Values	Description
httptype=<string>	singlepart, multipart	Choose streaming method. Default value is defined by the parameter Audio.A#.HttpMessageType
camera=<string>	1, 2, 3, 4, quad	Which audio configuration that should apply. camera=1 will get the audio configuration from the parameter group Audio.A0, the configuration connected to camera 1. camera=2 will get the audio configuration from the parameter group Audio.A1, the configuration connected to camera 2 etc. camera=quad will get the audio configuration from the parameter group Audio.A4, the configuration connected to the quad stream.

Example: Request a singlepart audio stream

```
http://myserver/axis-cgi/audio/receive.cgi?httptype=singlepart
```

3.3 Singlepart audio data response

When an audio stream is requested/transmitted, the server returns/receives a continuous flow of audio packets. The content type is only set at the beginning of the connection. When the connection is up and running the audio packets will come right after another without any extra information between the packets. The message body contains a block of binary data. Each block of coded audio data is 240 byte.

Return:

```
HTTP/1.0 200 OK\r\n
Content-Type: <audio MIME>\r\n
\r\n
<Audio data>
```

Example: Singlepart Audio data encoded with G.711 μ -law.

```
HTTP/1.0 200 OK\r\n
Content-Type: audio/basic\r\n
\r\n
<Audio data>
<Audio data>
```

```
<Audio data>
.
.
.
```

3.4 Multipart audio data response

When an audio stream is requested/transmitted, the server returns/receives a continuous flow of audio packets. The content type is "multipart/x-mixed-replace" and each audio packet ends with a boundary string. The message body contains a block of binary data. Each block of coded audio data is 240 byte.

Return:

```
HTTP/1.0 200 OK\r\n
Content-Type: multipart/x-mixed-replace;boundary=--<boundary>\r\n
\r\n
--<boundary>\r\n
<audio>
```

where the proposed <boundary> is:

myboundary

and the <audio> field is

```
Content-Type: <audio MIME>\r\n
\r\n
<Audio data>\r\n
--<boundary>\r\n
<audio>
```

Example: Multipart Audio data encoded with G.726 32kbit/s (G.721).

```
HTTP/1.0 200 OK\r\n
Content-Type: multipart/x-mixed-replace;boundary=myboundary\r\n
\r\n
--myboundary\r\n
Content-Type: audio/32KADPCM\r\n
\r\n
<Audio data>\r\n
--myboundary\r\n
Content-Type: audio/32KADPCM\r\n
\r\n
<Audio data>\r\n
--myboundary\r\n
Content-Type: audio/32KADPCM\r\n
\r\n
<Audio data>\r\n
--myboundary\r\n
Content-Type: audio/32KADPCM\r\n
\r\n
<Audio data>\r\n
--myboundary\r\n
.
.
.
```

3.5 Audio data transmit

Transmit a Singlepart/Multipart Audio data stream.

Method: POST

Syntax:

```
http://<servername>/axis-cgi/audio/transmit.cgi
```

There are no valid parameters and values.

Example: Singlepart/Multipart Audio data

```
http://myserver/axis-cgi/audio/transmit.cgi
```

4. AUDIO PARAMETERS

4.1 Audio

Description: Common audio parameters used for all audio configurations.

Configuration file: /etc/sysconfig/audio.conf

Security level (create/delete): 7/7

[Audio]

Parameter name	Default value	Valid values	Security level (get/set)	Description
DuplexMode	full	full, half, get, post, off	1/4	How the audio should be transferred. full = Full duplex - simultaneous two-way audio. Transmit and receive audio at the same time. half = Half duplex - non simultaneous two-way audio. Audio only allowed in one direction at a time. get = Simplex. Audio from the server. post = Simplex. Audio to the server. off = no audio sent or received.
ForwardOnPost	no	yes, no	4/4	The audio transferred from a client to the server, is also transferred to the other clients, i.e. the server is relaying audio. Note: This only works in half-duplex mode.
MaxListeners	10	1 ... 10	1/4	Max number of simultaneous clients (does not affect multicast delivery).

NbrOfConfigs	Product dependent	An unsigned integer	4/7	Number of audio configurations. The default value is product dependent. 1 for products with one image source and 5 for products using four image sources (4 inputs + Quad).
--------------	-------------------	---------------------	-----	--

4.2 Audio.A#

Description: Audio configuration. The name of the audio configuration, which audio source it is connected to and whether it should be enabled or not.

Configuration file: /etc/sysconfig/audio.conf

Security level (create/delete): 7/7

[Audio.A#] *

Parameter name	Default value	Valid values	Security level (get/set)	Description
Enabled	no	yes, no	1/4	Enable/disable this audio configuration.
HTTPMessageType	singlepart	singlepart, multipart	1/4	How the audio should be streamed. Some proxies require multipart streaming.
Name		A string	4/4	The name of the configuration.
Source	0	An integer	4/4	Which AudioSource.A# this Audio.A# configuration is connected to.

* **Note:** The # is replaced with a group number starting from zero, e.g. Audio.A0

4.3 AudioSource

Description: The number of audio sources (audio inputs/chips).

Configuration file: /etc/sysconfig/audio_source.conf

Security level (create/delete): 7/7

[AudioSource]

Parameter name	Default value	Valid values	Security level (get/set)	Description
NbrOfSources	1	An unsigned integer	0/7	Number of audio sources (audio input/chips).

4.4 AudioSource.A#

Description: Parameters for each audio source (audio input/chip).

Configuration file: /etc/sysconfig/audio_source.conf

Security level (create/delete): 7/7

[AudioSource.A#] *

Parameter name	Default value	Valid values	Security level (get/set)	Description
AcousticEchoCancellingEnabled	yes	yes, no	4/4	Suppress audio from speakers that go to the microphone. This parameter works on outgoing sound from the server.
AudioEncoding	Hardware dependent	g711, g721, g723 ¹	1/4	The audio codec to use.
InputGain	0	mute, -60, -57, -54 ... -3, 0, 3, 6 ... 21, 24, 27, 30	4/4	Gain setting in dB for sound received from client.
InputPreGain	high	low, high	4/4	Pre amplifier. Ability to compensate for different signal levels.
InputType	Hardware dependent	internal, mic, ₂ line ²	4/4	Where the audio shall be captured from and on which level.
LineEchoCancellingEnabled	no	yes, no	4/4	Suppress audio from speakers go to the microphone. This parameter works on incoming sound from the client.
MicrophonePower	off	on, off	4/4	Enable/disable power on audio input connector.
Name	Audio	A string	4/4	Name of the audio source (audio input/chip).
NoiseCancellingAttenuation	0	0, 8, 13, 17	4/4	Set the level of noise reduction.
NoiseCancellingIOGain	18	0, 6, 12, 18	4/4	Set the noise threshold value.

OutputGain	0	mute, -60, -57, -54 ... -3, 0, 3, 6 ... 21, 24, 27, 30	4/4	Gain setting in dB for sound transmitted to client(s).
SlopeFilterEnabled	yes	yes, no	4/4	If yes a filter reduces noise by filtering out all frequencies below 100 Hz and all above 3500 Hz.

* **Note:** The # is replaced with a group number starting from zero, e.g. AudioSource.A0

¹ Product-dependent. Check the corresponding Property parameter.

² Product-dependent.

4.5 Properties.Audio

Description: The Properties parameters contain information about product dependent functionality. The Properties parameters are product dependent. If a parameter does not exist, the functionality is not supported.

Configuration file: /usr/etc/param/par_prop_audio.conf

Security level (create/delete): 7/7

[Properties.Audio]

Parameter name	Default value	Valid values	Security level (get/set)	Description
Audio		yes, no	0/7	The product has audio support.
Format		A string	0/7	The supported formats separated by commas, e.g. g711,g721,g723.

5. AXIS ACTIVEX COMPONENTS

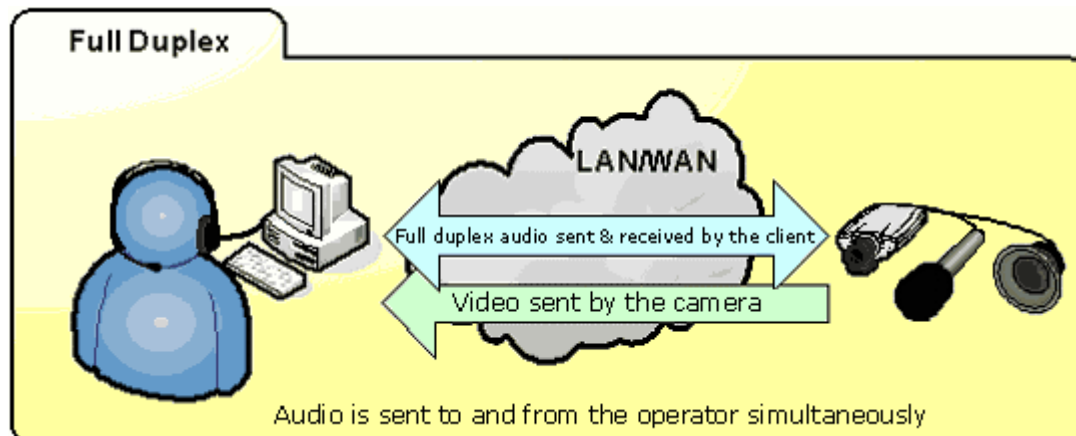
The default audio codec in Axis products using firmware version 4.xx is G711. This audio codec is supported by the **AXIS Media Control** but not by the AXIS Camera Control. The audio codec used can however be changed to G721 or G723 that are supported by the AXIS Camera Control. Change the audio encoding by setting the parameter AudioSource.A#.AudioEncoding to either g721 or g723.

6. AUDIO SETTINGS

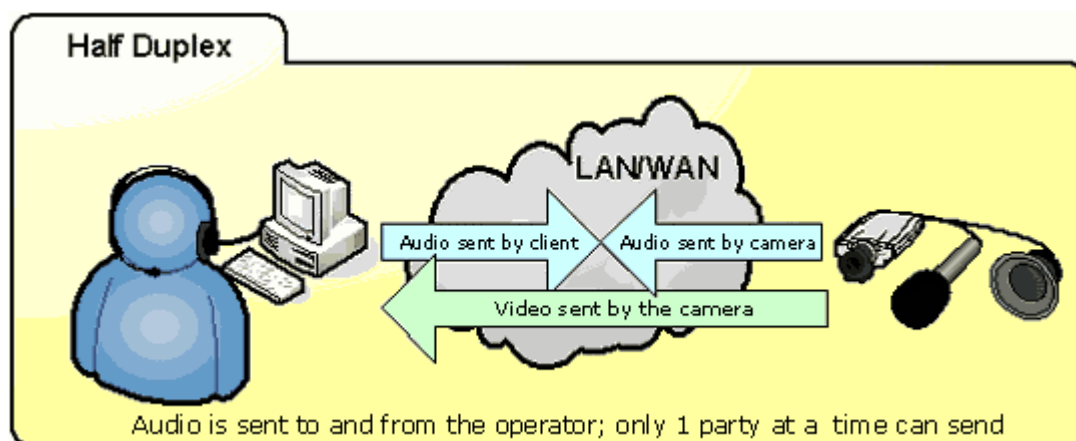
6.1 Audio modes

Two-ways audio mode:

Full Duplex – simultaneous audio in both directions. This mode requires a full-duplex sound card on the computer. If the available bandwidth is 0.2Mbit/s or less, it is recommended to use Half Duplex mode.

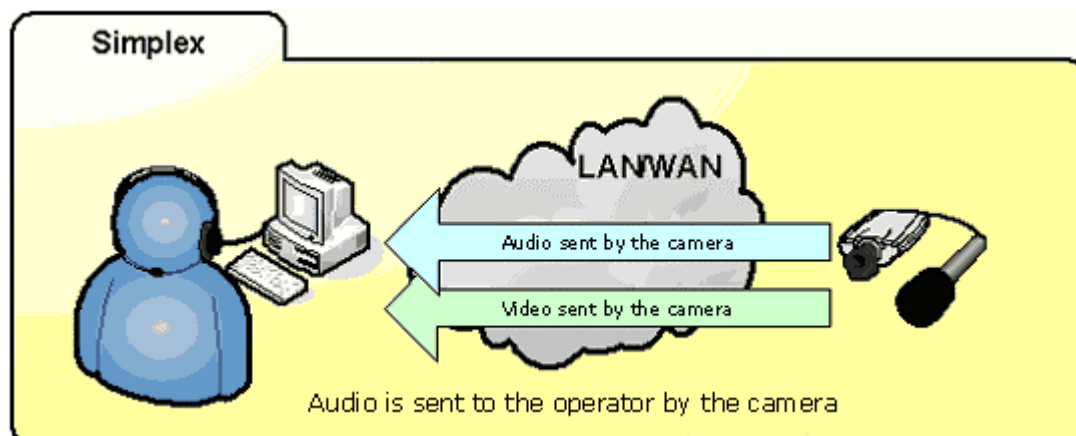


Half Duplex – one direction at a time for bandwidth efficiency.

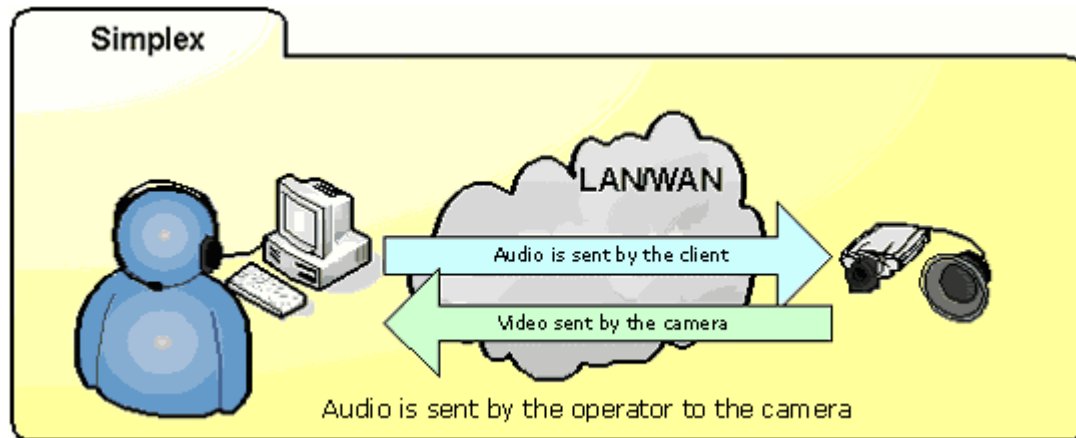


One-way audio modes:

Simplex – get mode. Audio is sent from the camera or video server to the client application.



Simplex – post mode. Audio is sent from the client application to the camera or video server.



6.2 Max number of clients

The parameter `Audio.MaxListeners` can restrict the number of clients accessing the camera. This is useful if bandwidth is limited. When using full duplex mode, up to 10 clients can receive video and audio from the camera, but only one client can transmit audio. If the application requires audio to be transmitted and received from more than one PC client, the half duplex mode should be used instead.

6.3 Broadcast the audio from the active client

Setting the parameter `Audio.ForwardOnPost` to yes will send the sound from the active client to all other connected clients. Note that this option requires using half duplex mode.

6.4 Audio input

The camera can pick up audio from its built-in internal microphone or from an external audio source connected to the audio input connector. The external audio source can either be a microphone or a line level source e.g. a CD player or audio mixer. Which to use can be set in the parameter `Audio.InputType`.

The sensitivity for the audio input source, e.g. the internal or external microphone, can be set to high or low, using the parameter `AudioSource.A#.InputPreGain`, depending on the sound level in the monitored area, e.g. select low if the camera is installed in a noisy environment.

The parameter `AudioSource.A#.InputGain` can be used to control the volume of the audio. Note however that using a low sensitivity for the audio input source (`AudioSource.A#.InputPreGain=low`) should never be compensated by a high input gain, as this may cause unwanted noise.

Power for the external microphone can be provided by enabling the parameter `AudioSource.A#.MicrophonePower`. This parameter must be set to on if using a small electret condenser microphone such as a clip-on microphone or a PC microphone. If using a high-impedance dynamic microphone the `MicrophonePower` parameter should be set to off. If using a professional microphone requiring 48V phantom power, an external power supply is needed with a balanced-unbalanced converter (audio transformer) in between.

6.5 Audio output

The audio output can be connected to a public address (PA) system or an active speaker with a built-in amplifier. A pair of headphones can also be attached. A stereo connector must be used for the audio out.

The volume of an active speaker attached to the camera can be controlled using the parameter `AudioSource.A#.OutputGain`.

6.6 Audio quality settings

Parameters are available to set the audio quality, to tune the internal or external microphone connected to the camera to suit your environment.

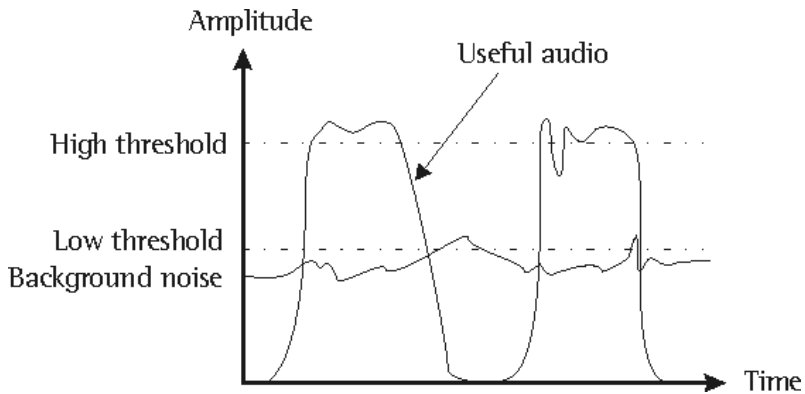
Speech filter

Setting the parameter `AudioSource.A#.SlopeFilterEnabled` to yes, reduces noise by filtering out all frequencies below 100 Hz and all above 3500 Hz. This improves the sound quality when the microphone is placed close to the person talking and can also help reduce background noise.

Echo cancellation

Setting the parameter `AudioSource.A#.AcousticEchoCancellingEnabled` to yes, reduce acoustic echoing that can occur if audio from the speakers is captured by the microphone.

Noise cancellation



Noise cancelling is a way of reducing the background noise when there is no useful audio. A typical application is that the camera is set up in a noisy environment, and you are only interested in hearing sound e.g. when a person is talking close to the microphone.

There are two parameters available to optimize this function:

- Noise canceller threshold level (`AudioSource.A#.NoiseCancellingIOGain`)
- Noise canceller attenuation (`AudioSource.A#.NoiseCancellingAttenuation`)

When the incoming sound is louder than the threshold, it will pass without any changes. When the incoming sound is weaker than the threshold, it will be reduced by a certain attenuation factor (measured in dB).

The threshold value should be set higher than the background noise, but lower than the useful audio. A lower threshold will accept most of the audio, only the weakest background noise will be reduced. A

higher threshold will make the noise canceller act on even stronger background noise.

The noise canceller attenuation produces maximum noise reduction, when set to 17dB of attenuation. If the Noise canceller attenuation is set to 0dB, no noise cancellation is performed.

Note: At the maximum noise canceller threshold value, there is a risk of reducing useful audio. There is a trade-off between noise cancelling and sound quality, increasing the noise attenuation deteriorates sound quality, and decreasing the noise attenuation improves sound quality.

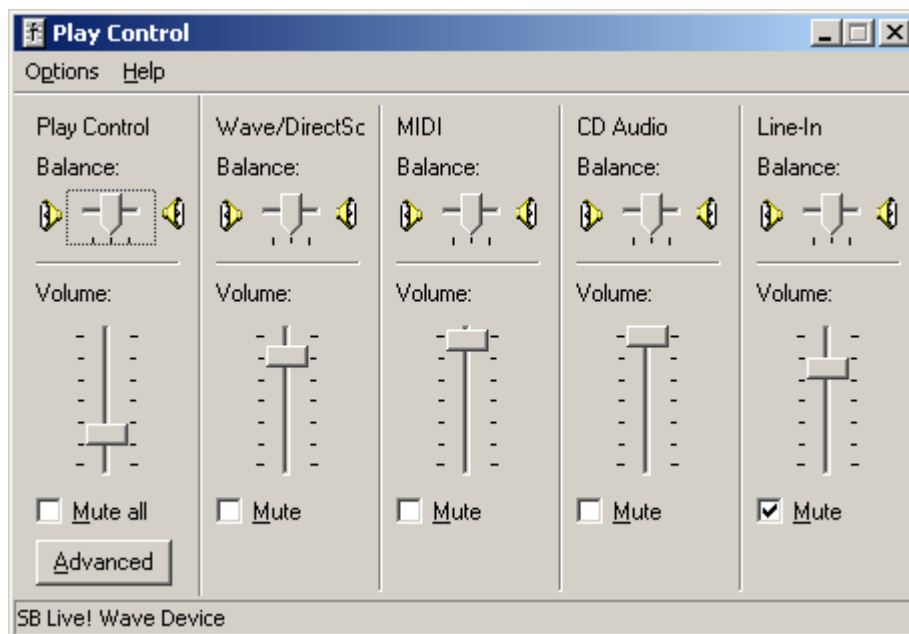
Optimize the threshold and the attenuation by listening and changing the levels.

6. CLIENT COMPUTER SETTINGS FOR SENDING AUDIO TO THE CAMERA

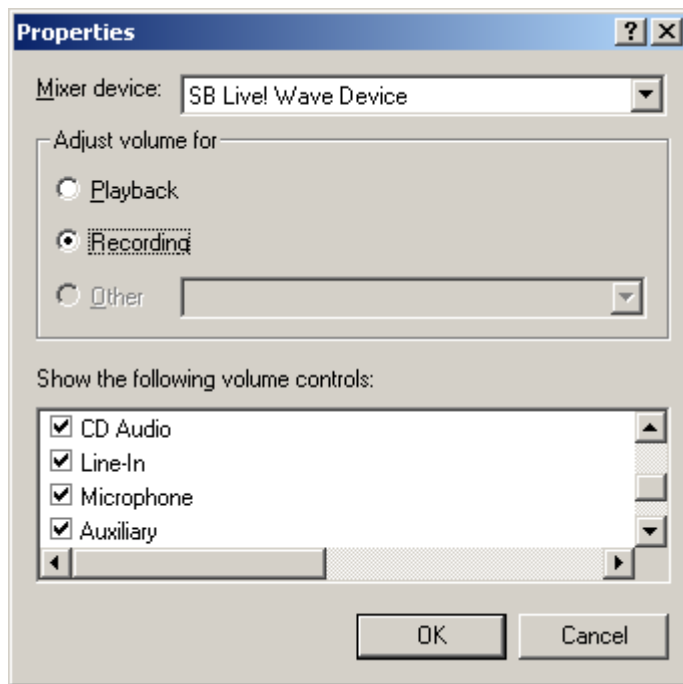
It is possible to send audio captured from a microphone, from a file stored on your computer or even a mix of both, to the camera.

In Windows, double-click on the speaker icon or select Start -> Accessories -> Entertainment -> Volume control.

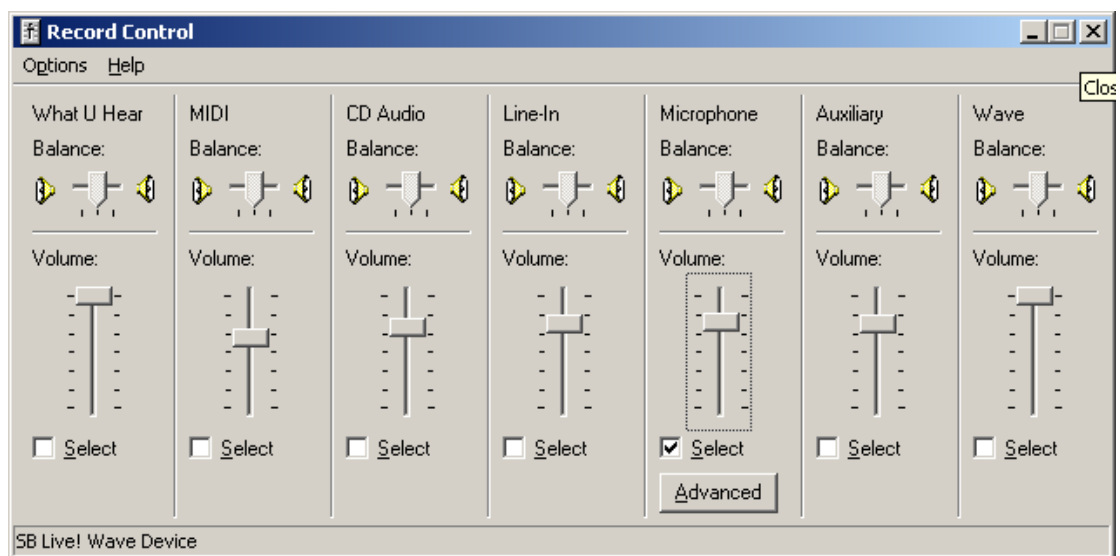
A Play Control menu will appear:



This menu controls sound produced by the computer. To configure recordings or send audio to the camera, click on Options -> Properties. This menu will appear:



Select Recording. To display all the relevant volume controls, check the corresponding check boxes in the list, especially "Microphone". Click OK.



Now you can choose to play sound from the computer to the camera by selecting one of the audio options. Microphone is normally selected but you can also play a sound file by selecting "Wave". In this mode any sound from a media player will be transmitted over the network to the camera.

To play a background sound mixed with the microphone signal, select "What U Hear". This provides a mix of all available sound. The corresponding volumes can be set individually.

