Perceptual *ACE*^{HD}

Real Time Video Encode Server

User Guide 1.1

DISCLAIMER AND COPYRIGHT NOTICE

TABLE OF CONTENTS

1. INTRODUCTION

ATC Labs has developed an efficient perceptual based real-time video encode server, **Perceptual ACE^{HD}.** Our latest innovation in video compression saves bandwidth costs significantly. Nearly 40% bandwidth efficiency can be achieved using ATC Labs perceptual analysis model. Perceptual ACE features state-of-the-art H.264 codec, enhanced using ATC Labs propriety Perceptual Video Analysis Platform; PVAM technology to deliver an excellent, crispy clear video quality at an impressively lower data rate. Perceptual ACE^{HD} uses an optimized H.264 codec to deliver high performance, high quality video including HD.

2. HARDWARE DESCRIPTION

The hardware configuration is chosen to support real-time operation of the Perceptual ACE^{HD} software for video sources in different formats including HD video. The basic hardware configuration is as follows:

- 2-socket tower based or 1U rack mount Dell Server with Eight-Core Intel Xeon CPU (Dual Quad-Core Xeon Processors), 2.66 GHz clock rate
- At least 8 GBytes RAM Memory
- Networking with dual or more Ethernet ports
- Windows XP operating system
- 60 GB Primary disk (additional secondary disks may be added)
- Blackmagicdesign DeckLink HD Extreme card with SDI video/audio input
- DVD to SDI adapter

3. SOFTWARE APPLICATION AND GUI

ATC Labs Perceptual ACE^{HD} accepts input from either a video device in real time or a raw video file and performs the video encoding and/or analysis in real time. The video encoding is either using the built in ATC Labs Perceptual H.264 encoder or alternatively a proprietary video encoder may be employed (by including a DLL with a specified API). In the case of input from a video card the software engine is fully tested with Blackmagic DeckLink Extreme HD card, but it is expected that it will continue to work with any DirectShow enabled video device. The supported video formats are described later in this document. It incorporates built in capability for YUV format conversion (e.g. 422 or 444 to 420 Planar) and frame rate conversion (e.g. 29.9 to 23.976 fps). The real time module also allows for optional capture and encoding of audio in the AAC format. The output of the real-time engine may be saved into compressed audio, video, and, data (in the case of video analysis) files. Furthermore, the compressed video and audio streams may be embedded into a transport format (e.g. MPEG2 TS) and optionally broadcast to a multicast connection. In the case of multicast data of H.264 video the streams may be monitor on another networked computer using VLC player. Various audio and video parameters may be specified using the GUI for desired results.

3.1   Supported Input Video Formats

The ATC Labs PVAM accepts inputs in digital raw YUV 422, 444 and, 420Planar formats. For the purpose of analysis and encoding the output is internally converted to the 420Planar format. It accepts most commonly used picture sizes such as CIF (352x288), 4 CIF (704 × 576), 16 CIF (1408 × 1152), DVD (NTSC 720 x480 or PAL 720 x 576), HDTV/Blue-ray-720p (1280 x 720), HDTV/Blue-ray-1080p (1920 x 1080), VGA (640 x 480), XGA (1024 x 768), etc. Furthermore support for interlaced video 1080i/720i is also provided using a built in pull-down conversion module.

4. PERCEPTUAL ACE^HD INTERFACE OVERVIEW

Perceptual ACE^HD GUI is used at the transmitting end to interface with the video device(s) in real time or raw video file and perform the video encoding; the output may be saved as a file (with a variety of supported container formats) or sent out over IP network as unicast/multicast stream. VLC player is recommended as the preferred software client to receive and display the output.

4.1 Starting the Perceptual ACE^HD GUI

To launch the application, go to **Start > All Programs > Perceptual Ace > Perceptual Ace 1.0** or locate the Perceptual Ace 1.0 shortcut from the desktop.

4.2   VIDEO PARAMETERS
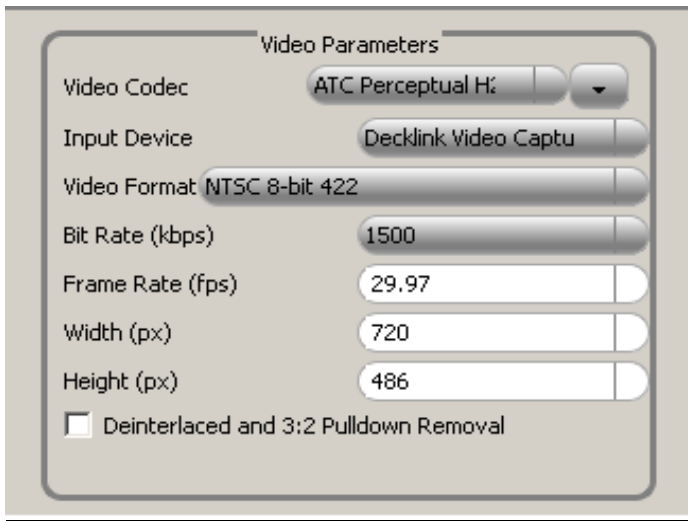
4.2.1   Video Codec

Click on the drop down menu to choose the desired video codec; ATC Labs perceptual H.264 or any other alternative proprietary video encoder (by including a DLL with a specified API) to process the file.
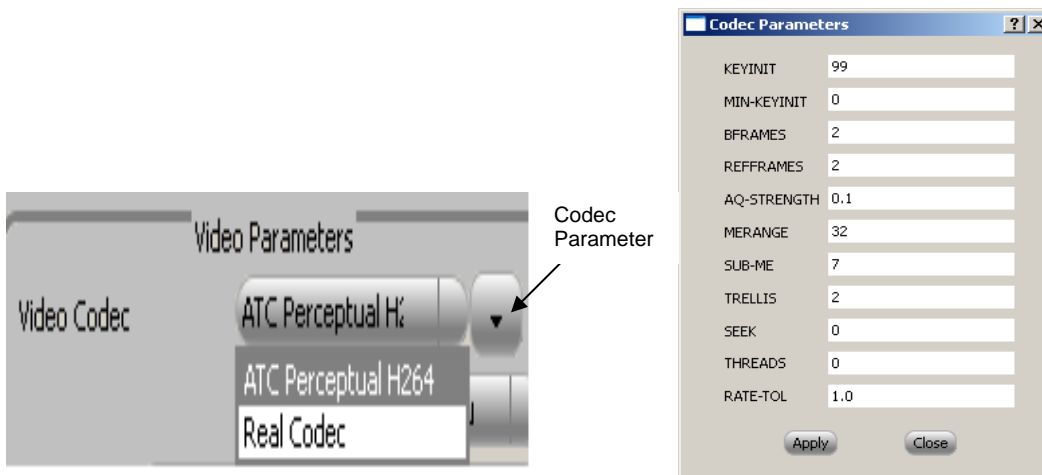
Codec Parameter

4.2.2   Codec Parameters

Click on the drop down menu as shown above to display the codec parameters. Currently this displays the codec "ini" settings of the Perceptual ACE. The codec parameters can be specified for alternative proprietary video codec using this option.

4.2.3 Input Device

The input can be either file based or from a video card. Select the input device using the drop down menu. The software automatically detects and lists the set of video capture cards installed on your computer.
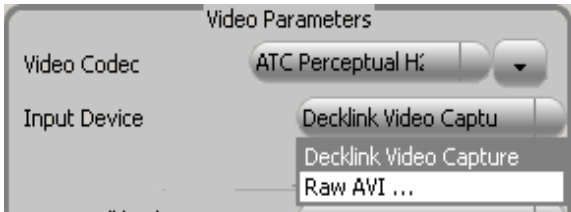
4.2.4 Video Format for a Raw AVI file

If the input is a raw avi file, color format needs to be specified. Once the raw avi file is opened, the color format can be selected using the **Video Format** option.
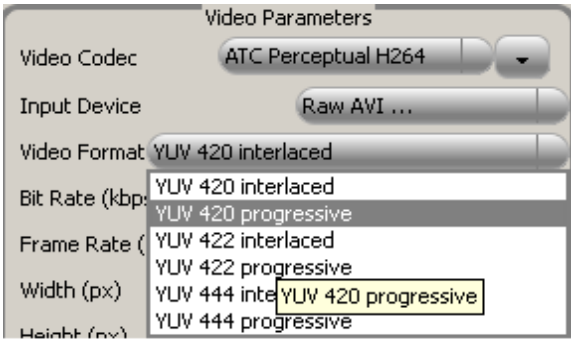
4.2.5 Video Format for a Device

The software automatically detects and lists the video formats available on the input video device. Select the video format using the drop down menu. This video format selection should complement the mode selection on the external device (480i, 720p, 1080i). Currently for the Blue Ray the output fed through the SDI converter, three of the formats are relevant.

480i	NTSC 8bit 422
720p	HD720 59.94p 8 bit 422
1080i	HD1080 29.97PsF 8 bit 422

7
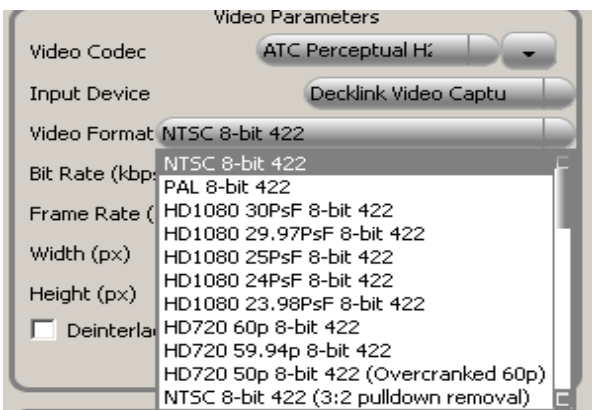
4.2.6   Bitrate (kbps)

Select the desired bit rate from the drop down menu to process the video. The allowed bit rate will depend on the mode selected on the external device (480i, 480p, 720p, 1080i) and the desired output size.

Bit Rate (kbps)	1500
Frame Rate (fps)	600
	700
Width (px)	800
	900
Height (px)	1000
☐ Deinterlaced and 3:2 Pull	1500
	2000
	2500
	3000
Audio Par	3500
	4000
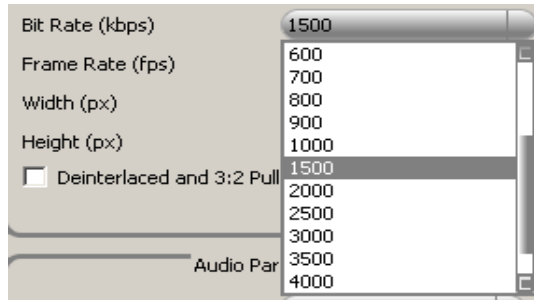
4.2.7   Framerate (fps)

Select the desired frame rate from the drop down menu to process the video. The available frame rates are as follows 10fps, 15fps, 20fps, 24fps, 30fps.

Frame Rate (fps)	29.97
Width (px)	10.0
	15.0
Height (px)	20.0
	24.0
☐ Deinterlaced and 3:2 Pull	30.0
	29.97
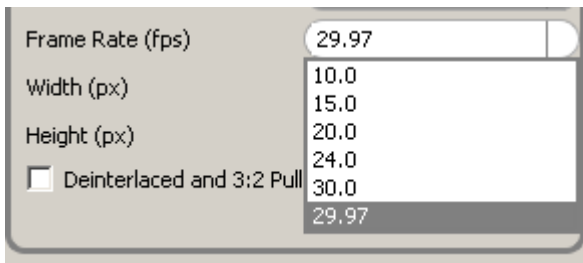
Default recommended output frame rate selection:

480i	720p	1080i
29.97fps	29.97fps	*23.98fps
*23.98fps	23.98fps	

*23.98fps can be selected after checking the Deinterlaced box on the GUI.

Frame Rate (fps)	23.98
Width (px)	720
Height (px)	486
☑ Deinterlaced and 3:2 Pulldown Removal	

8
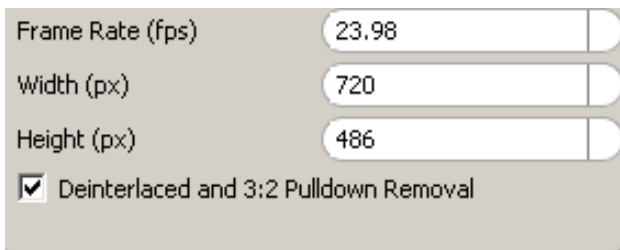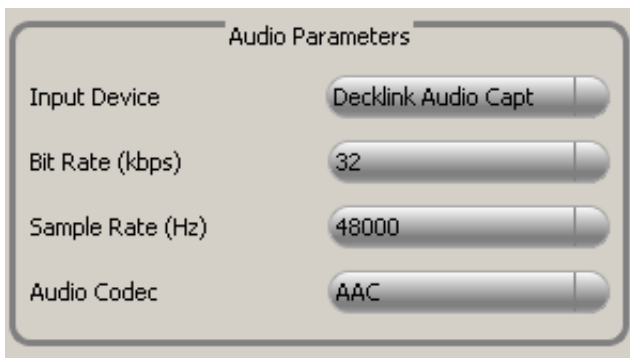
4.2.8   Width/Height

Select the desired width and height of the output.

Width (px)	720
Height (px)	486
☐ Deinterlaced and 3:2 Pulldown Removal	

4.3   AUDIO PARAMETERS

Audio Parameters

Input Device	Decklink Audio Capt
Bit Rate (kbps)	32
Sample Rate (Hz)	48000
Audio Codec	AAC

4.3.1   Input device

The software automatically detects and lists a set of audio capture cards installed on your computer. For the current set up, the audio input device is a Declink Audio Capture Card.

4.3.2   Bitrate (kbps)

Select the desired audio bitrate using this option. The bitrate selection options are as follows 32kbps, 48kbps, 64kbps, 96kbps, 128kbps.

4.3.2   Sample Rate (Hz)

The sample rate is device specific. The software automatically detects and displays the sample rate.

4.3.3   Audio Codec

Choose the audio codec from the list. The available codecs are AAC and MP2.

4.4 OUTPUT
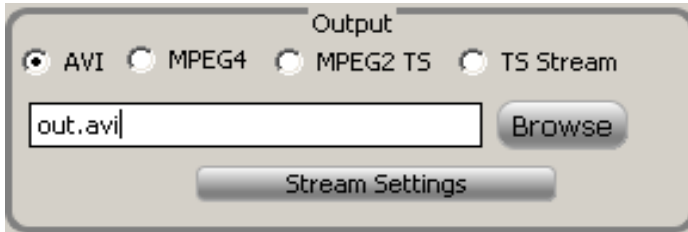
The output of the processed file could be saved in container format such as avi/TS/MPEG4. The TS container format can be streamed using a multicast/unicast socket. Select the TS Stream option on the Output menu selection. With this option all the IP address, port information can be specified using the Stream Settings.

4.4.1 Unicast

To stream to a single computer (Unicast); enter the client's/receiver's actual IP address in the 0.0.0.0 - 223.255.255.255 range; using the port 8080 is recommended for UDP unicast.
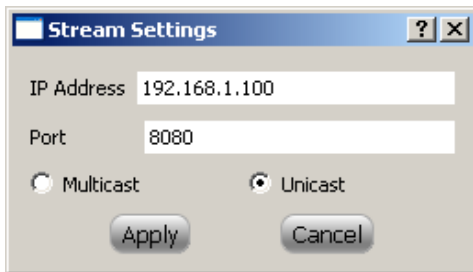
4.4.2 Multicast

To stream to multiple computers using multicast; enter the actual IP address of the multicast group in the 224.0.0.0 to 239.255.255.255 range (example: For creating an intranet multicast group the default IP of 238.38.38.38 typically works well). Port 4000 is recommended as an initial setting for UDP multicast.
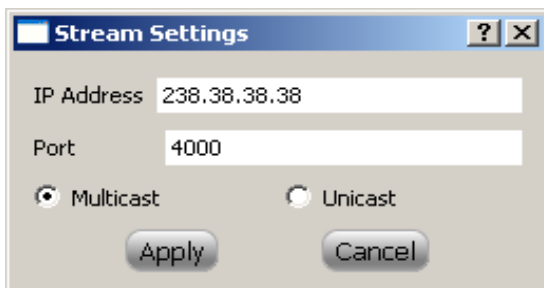
5. VLC SET UP FOR PLAYBACK & STREAMING

5.1 AVI AND MP4 FILE PLAYBACK

For playback of AVI and MP4 files VLC player needs to be set up to use the default demuxer. This is explained here (for the play back of TS files setup is similar to the streaming mode setting as described later in this section). On the VLC media player menu choose **Tools > Preference,** select **ALL in show settings.**

5.1.1 Tools > Preference > Input/Codecs > Demuxer > Demuxer Module: DEFAULT
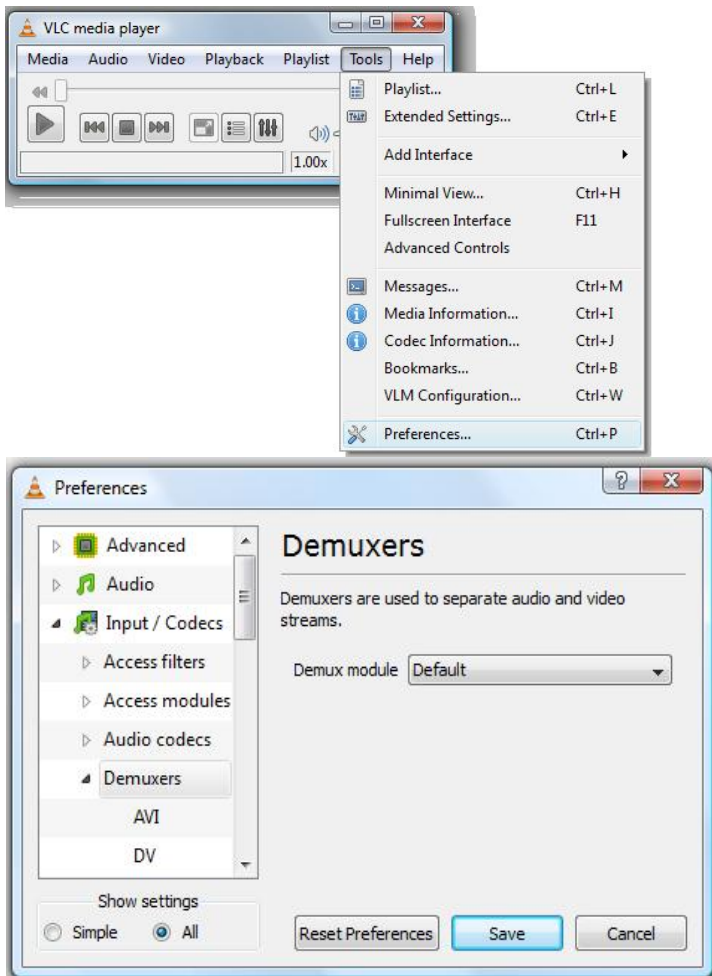
5.1.2 Tools > Preference > Input/Codecs > Access Module > File > Caching value: 800ms
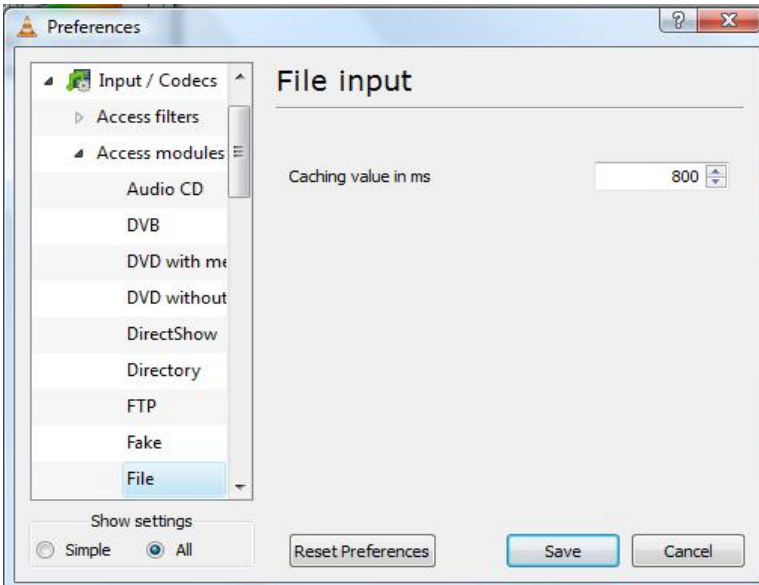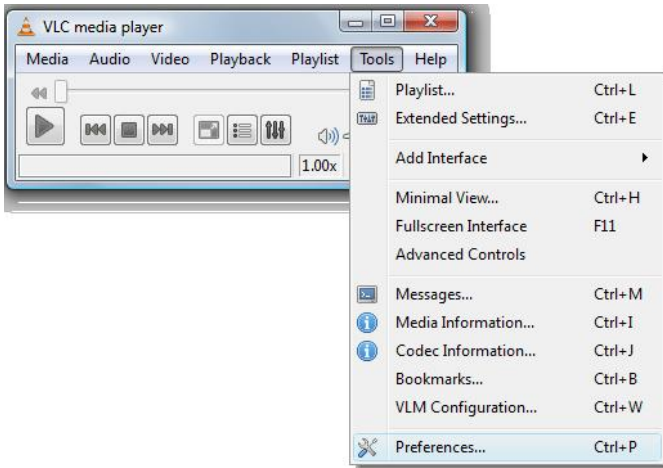
5.2 STREAMING

VLC Demuxer module settings for Streaming and TS File Playback

On the VLC media player menu choose **Tools > Preference,** select **ALL in show settings.**

Tools > Preference > Input/Codecs > Demuxer > Demuxer Module: FFmpegdemuxer
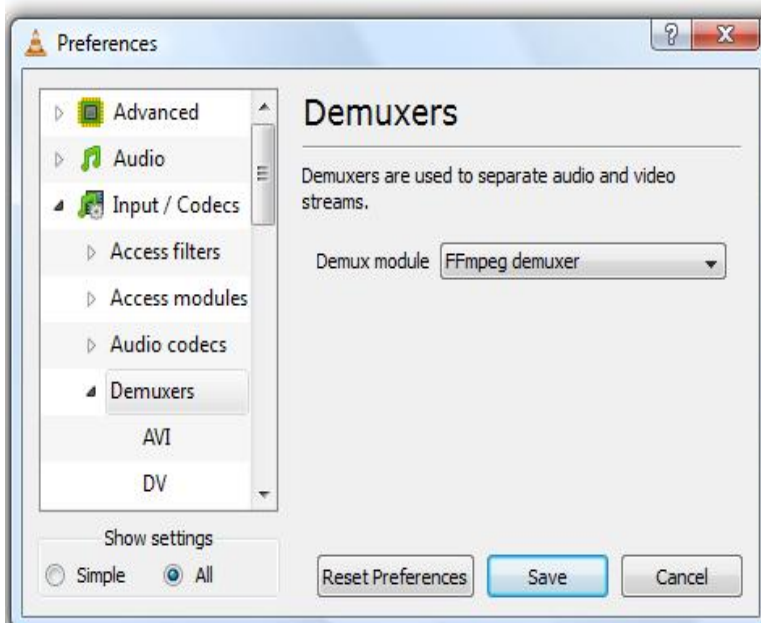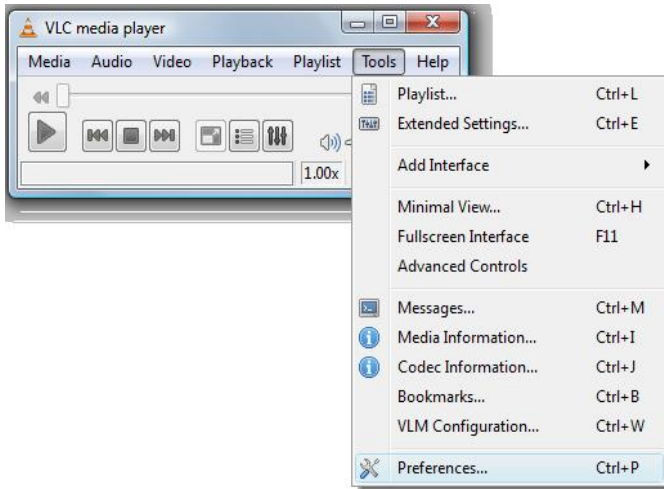
5.2.1 Stream from network option

- UDP Unicast: This applies to the case when on Perceptual ACE^HD Unicast Stream option to a single computer is chosen. Enter the client's/receiver's IP address. And choose port setting matching the Perceptual ACE^HD setup (e.g. 8080).

- UDP Multicast: This applies to the case when on Perceptual ACE^HD Multicast Stream option to multiple computers is chosen. Enter the IP address of the multicast group. And choose port setting matching the Perceptual ACE^HD setup (e.g. 4000).

On the VLC media player choose **Media > Open Network** or choose **Media > Streaming > Network Tab** and select the **Network Protocol > UDP.** Enter the IP Address based on Unicast or Multicast option as described above and set the **Port > 8080(Unicast)/4000(Multicast)** matching the Perceptual ACE^HD set up.

Also check the Show More Options box and specify the caching value to **2000 ms**. Finally press stream button next to the caching value.
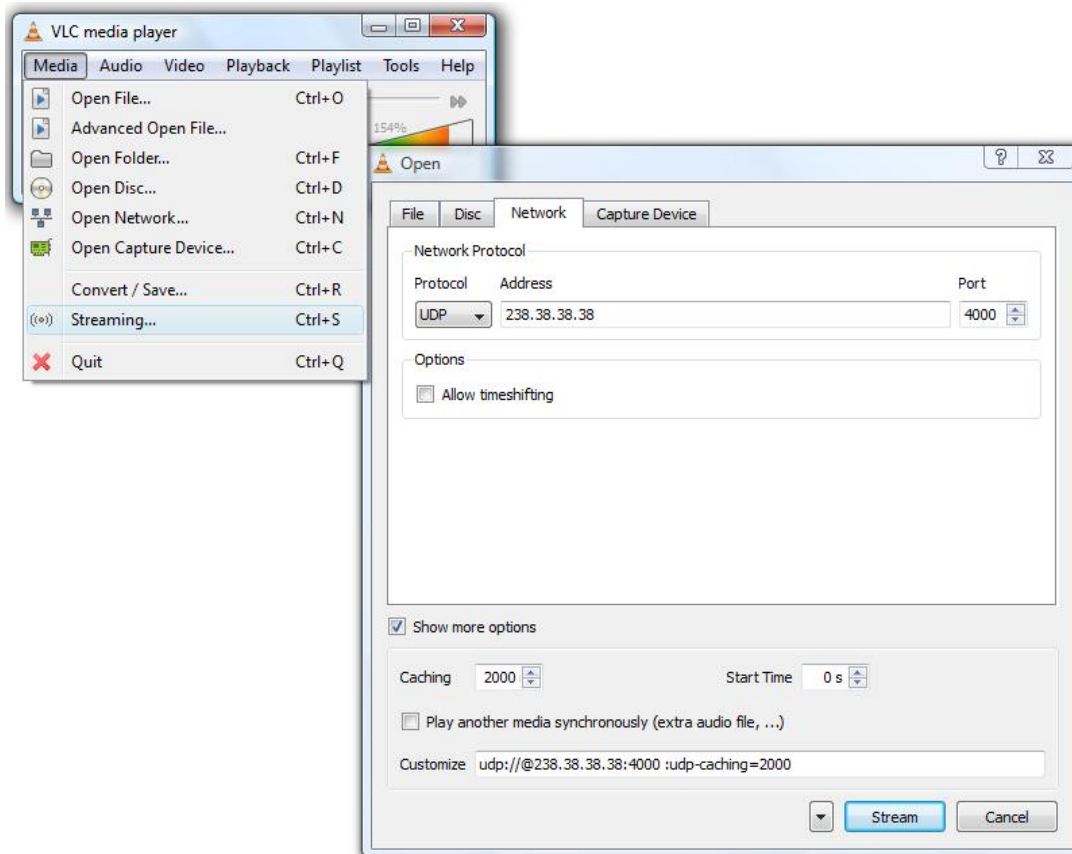
6. CONTACT INFORMATION

Corporate Head Office, USA

ATC LABS

Audio Technologies and Codecs, Inc.

105 Lock Street

Suite 411

Newark, NJ 07103

Tel: (973) 624-1116

Fax: (973) 624-1118

Customer Support

Email: support@atc-labs.com

Tel: (973) 796-2840

General Information

info@atc-labs.com

Software and Technology Development Center, Noida, India

Audio Technologies and Codecs (India) Pvt. Ltd.

A-87 Ground Floor

Sector 4

Noida-201301 (U.P), India

Tel: +91-120-2519611

Fax: +91-120-2519610

APPENDIX

7. HARDWARE SET UP

7.1 Example of Hardware Configuration Using ONKYO Blu Ray DVD Player

a) Dell PowerEdge 2900 III server with Dual Quad Core Intel Xeon, 3.16 GHz clock rate
b) 16 GB RAM Memory
c) Networking with dual or more Ethernet ports
d) Windows 2003 operating system
e) 1TByte primary disk and 1 TByte secondary disk
f) Pre-installed Blackmagic Design DeckLink HD Extreme card with SDI video/audio input
g) DVD to SDI adapter
h) Pre-installed Matrox M9120 Plus LP PCIe X1 graphic card

i) Dell PowerEdge 2900 III server with Dual Quad Core Intel Xeon, 3.16 GHz clock rate
j) 16 GB RAM Memory
k) Networking with dual or more Ethernet ports
l) Windows 2003 operating system
m) 1TByte primary disk and 1 TByte secondary disk
n) Pre-installed Blackmagic Design DeckLink HD Extreme card with SDI video/audio input
o) DVD to SDI adapter
p) Pre-installed Matrox M9120 Plus LP PCIe X1 graphic card
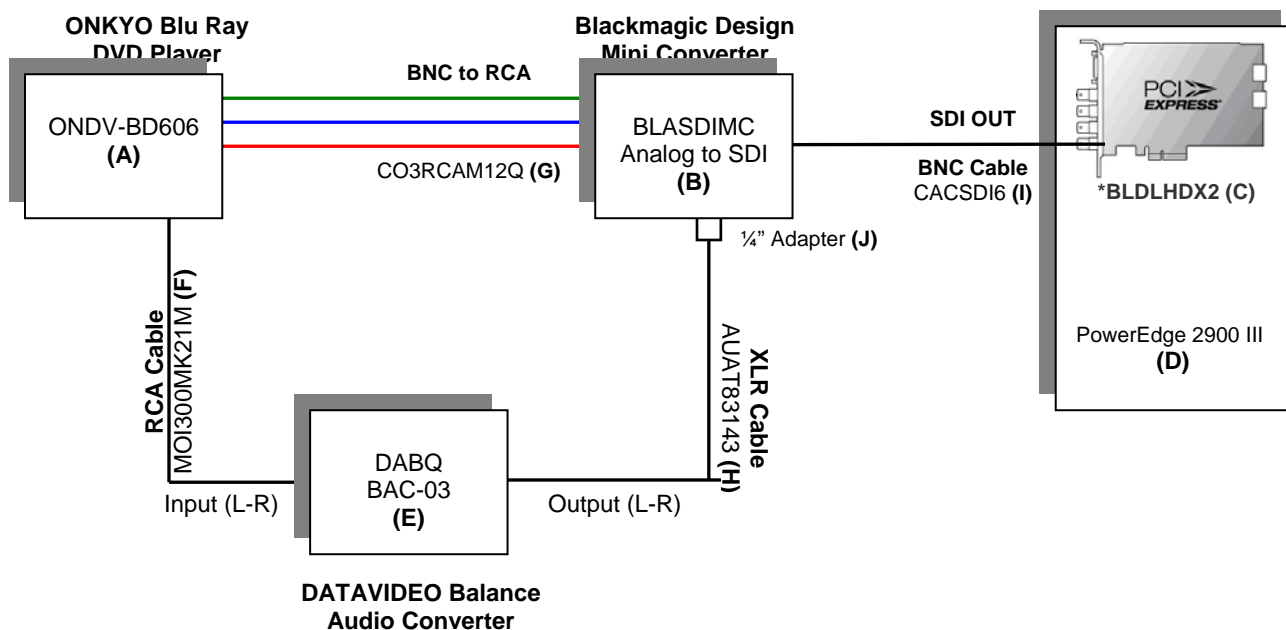
Parts List and description

PART NUMBER	PART NAME	DESCRIPTION
A	ONDV-BD606	ONKYO BLU-RAY Disc Player
B	BLASDIMC	Black Magic Analog to SDI Mini Converter
C	BLDLHDX2	*Pre-installed Black Magic DecLink HD X-Treme-2
D	Dell Server	PowerEdge 2900 III
E	DABQ BAC-03	DATAVIDEO Balance-Balanced Audio Converter
F	MOI300MK21M	Monster Cable Interlink 300 2 RCA/2 RCA Cable
G	CO3RCAM12Q	Communications Specialties 3XRCA M to 3XBNC
H	AUAT83143	Audio Technica 3' XLRF-XLRM Balanced MIC Cable
I	CACSDI6	Canare Serial DIGI Cable (SDI) BNC-BNC-6'
J	¼" Adapter	2 channels of professional balanced analog/AES/EBU audio with standard 1/4 inch jack connections.
	Matrox M9120	Pre-installed Matrox M Series Graphic Card

Note:

(A) ONDV-BD606
The initial video resolution for component output is set to 1080i. The component output setting can be changed using the setup menu on the DVD Remote Control:
1. In the stop mode, press SETUP.
2. Use up and down arrow to select the desired setting category: VIDEO
3. Select the setting option COMPONENT OUTPUT and set the desired video resolution.
** For more information on function set up, please refer to ONKYO BLU RAY DVD manual (page 34).

(B) BLASDIMC
Use the mini switches located on the side of the Black Magic Analog to SDI mini converter to change the settings. The switch information is as follows:

	OFF	ON
8	ANALOG Audio	AES/EBU Audio
7	7.5 IRE	0.0 IRE
6	SMPTE Levels	BETACAM Levels
5	Component	Composite or S-Video
4	Composite	S-Video
3		
2		
1	Processing OFF	Processing ON

8.   PVAM TECHNOLOGY
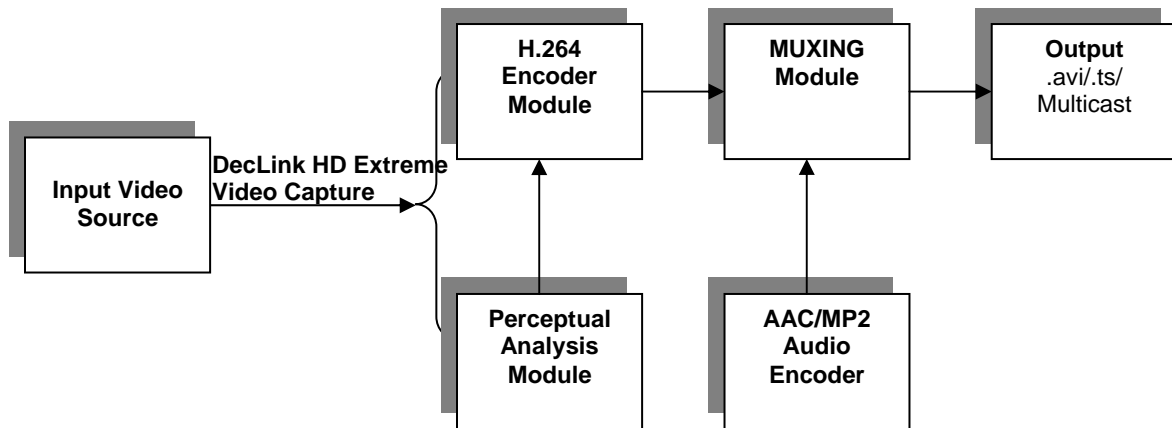
The ATC Labs Perceptual Video Analysis Platform (PVAM) performs detailed analysis of the video sequence and generates a set of perceptually meaningful measures including a Just Noticeable Distortion (JND) mask transparent coding. The perceptual analysis module comprises of the following:

8.1   Temporal Complexity Analysis

This module performs temporal analysis of the video sequence and generates a temporal complexity measure. The complexity measure is based on a number of parameters and provides a robust mechanism for detecting scene changes with very high rate of successful detection and very low probability of false positives. Furthermore, using appropriate thresholds the measure can also be used to flag scene transitions. This module also generates other useful information such a flag to indicate appropriate placement of I frames in a video codec desiring an adaptive GOP framing and is also suitable for use in bit allocation algorithms.

8.2   Texture Complexity Analysis

Texture describes the relative proportion of grain sizes in any region of the image. This module generates a unified texture measure for each macro-block (16 x16 regions) in each picture in the video sequence. The texture measure is normalized for intensity (i.e. is not affected by intensity) and usable as a single linear zed measure of texture complexity for the purpose of macro-block classification into low, mid, or, high texture classes. It has been rigorously verified through visual cross validation with a large database of real-time video data.

8.3   Edge Detection & Labeling Analysis

Edges play an important role in the perception of video as well as the masking of the quantization noise in a codec. Therefore the perceptual analysis engine incorporates a robust algorithm for edge detection and labeling with a pixel level accuracy. It has been carefully tuned to avoid detection of spurious edges.

8.4   Perceptual Just Noticeable Distortion (JND) Estimation

This module generates a JND mask in the form of maximum allowed distortion for each macro-block in each picture for transparent coding. The allowed distortions is computed as a PSNR number such that for normal viewing the distortion is expected to be visible to a trained viewer if the PSNR for the macro-block is any worse than the indicated transparency level. The JND or PSNR mask generation process uses a number of perceptual features as input and translates them into a target PSNR value using maps generated through extensive psycho-visual experimentations with a large video database consisting of idealized test patterns and real-life video sequences. Simplicity & generality are the other main features of the model which makes it implementable in real time applications. The perceptual features which are generated and provided to the PSNR mapping as input include:

- Texture measures which describe the relative proportion of grain sizes in any region of the image.
- Intensity measure, i.e., the average brightness level of an image area.
- Edge indicators, indicating a sudden discontinuity in the average brightness levels.
- Motion Activity, i.e. the relative speed of displacement of objects across video frames.

8.5   JND to Quantization Parameter Mapping Generator

This module generates an accurate mapping from the target JND/PSNR to a quantization parameter assuming H.264 quantization model. This so called inner loop functionality ensures that generated quantization parameter is optimum in the sense that it indicates the coarsest quantization step size for which the measured distortion is lower than input target PSNR. Once quantization parameter per macro-block is generated and it can take either the original or differential (i.e. the output of Inter & Intra predictors) video data as input.