Stereoscopic Player and Stereoscopic Multiplexer: a computer-based system for stereoscopic video playback and recording

Peter Wimmer[*]

Johannes Kepler University Linz, Altenberger Str. 69, 4040 Linz, Austria

ABSTRACT

Computer-based solutions are well suited for capturing and playing stereoscopic video content at high quality, because the constraints of TV standards need not be obeyed. We developed an application called *Stereoscopic Player*[1] to play stereoscopic videos. Based on Microsoft's DirectShow, it can handle all major video sources, codecs, stereoscopic layouts and viewing methods. Usability was an important goal in the application design; several features simplify handling of stereoscopic videos. To free the user from choosing the proper settings, we introduced the concept of stereoscopic information files and an Internet-based service for automatic configuration.

The *Stereoscopic Multiplexer*[2] is a solution to capture stereoscopic content within applications compatible with WDM capture drivers. It takes frames from two 'real' capture devices, synchronizes them and passes the resulting stereo pairs to the application. If the cameras support numerical control of camera parameters, Stereoscopic Multiplexer also synchronizes these parameters.

Supporting resolutions up to high definition and a wide range of devices, Stereoscopic Multiplexer and Stereoscopic Player are universal, affordable solutions for stereoscopic video recording and playback.

Keywords: stereoscopic video, capture driver, on-the-fly conversion, recording, playback, transmission, software

1. INTRODUCTION

Stereoscopic filmmaking is known for over hundred years and had several ups and downs in its history[3]. Stereoscopic video technology was also invented many years ago. But neither stereoscopic film nor video technology ever became a mainstream market, opposed to 2D film and video. The reasons are manifold: On the one hand, 3D technology is more complex than monoscopic technology. On the other hand, watching stereoscopic content either requires special glasses or limits the positions in front of the screen where a stereoscopic image can be perceived. The huge number of possible stereoscopic viewing methods, their advantages, disadvantages and dependencies on display technologies complicates setting up a viewing system. Whereas the ideal stereoscopic display still has to be invented, recording and playback of stereoscopic videos can already be dramatically simplified using a computer based approach. Computer based solutions bypass limitations of analog TV systems, namely limited color resolution (problematic for anaglyph viewing), low refresh rate (problematic for shutter glasses usage) and low resolution. Passive polarized projection also benefits from computer technology: The projectors can be connected to a dual head graphics card, eliminating the need for expensive, synchronized hardware players and allowing arbitrary resolutions beyond any TV standard.

This paper presents two software solutions which dramatically simplify handling stereoscopic video content. Stereoscopic Player is an easy to use video playback application, supporting a wide range of stereoscopic layouts and viewing methods. It was designed for standard personal computers and the Microsoft Windows operating system family. Based on Microsoft's popular DirectShow architecture, it can play almost all important file formats, 3D-DVDs, web streams and live content from capture devices. Stereoscopic Player was developed both for consumers and for professionals, making it an inexpensive alternative to specialized hardware solutions. Figure 1 shows its main window while playing an IMAX-3D-DVD.

---

[*] Email: pwimmer@gmx.at

Figure 1. Stereoscopic Player main window

Stereoscopic Multiplexer is a driver to capture stereoscopic video content from two distinct cameras. It can be used within any application compatible with WDM (Windows Driver Model) drivers. It takes video frames from two real capture devices, combines them to side-by-side format and passes the resulting stereoscopic image pairs to the application. Supported cameras include webcams, DV cameras and industrial Firewire cameras, but any other capture devices supported by Windows work as well.

When Stereoscopic Multiplexer and Stereoscopic Player are used together, live transmissions over the Internet are possible. Our test configuration runs Windows Media Server and Windows Media Encoder on the server side. The Windows Media Encoder must be configured to use Stereo Multiplexer as a virtual capture device; consequently, the stereoscopic video stream is broadcasted in side-by-side format. All other settings can be chosen equally to monoscopic live transmission. On the client PC, Stereoscopic Player receives the Windows Media stream and converts it to any supported viewing method.

2. CONCEPTS OF STEREOSCOPIC PLAYER

Usually, stereoscopic videos are encoded for a specific viewing method. If a 3D-DVD is targeted for fieldsequential playback using shutterglasses, the content is encoded in fieldsequential MPEG-2 format. The fieldsequential format uses the interlaced scanning system to store left and right images in odd and even scanlines or vice versa. If the DVD comes along with anaglyph glasses, the content is encoded in anaglyph format. The anaglyph format is also popular for content delivered via Internet and CD-ROMs. The main problem is the missing compatibility of the different formats: An anaglyph video cannot be watched in full color using shutterglasses or polarized projection, whereas fieldsequential videos cannot be watched with anaglyph glasses. Sometimes, DVDs contain both a fieldsequential and an anaglyph version to work around this problem. However, this approach requires double disk space, increases production costs and does not provide a solution for many other stereoscopic viewing methods available.

Universal playback capabilities can be provided by entirely decoupling encoding format and viewing method. The DVD, file or webstream is delivered in only one stereoscopic layout which provides the left and right perspectives in full color. The output format necessary for a certain viewing method is calculated on-the-fly. Possible distribution layouts include interlaced, side-by-side and over/under format. Anaglyph formats are not suitable since full color left and right perspectives cannot be recovered. Separate disks or files for the left and right views would be possible, too, but are more difficult to handle and can cause severe performance problems when located on the same disk. From the performance point of view, over/under would be the ideal format[4], whereas the side-by-side format has the advantage that it can be previewed without glasses using parallel or cross-eyed free viewing techniques. The interlaced format retains compatibility with fieldsequential shutterglasses systems, but wrong color interpolation in decoders often introduces ghosting and degrades image quality.

Stereoscopic Player implements on-the-fly conversion in software. It can play all file formats supported by DirectShow (for example AVI, MPEG and Windows Media) using any of the following stereoscopic layouts:

– Monoscopic
– Interlaced
– Over/under
– Side-by-side

Monoscopic files are not converted to 3D. Instead, the monoscopic layout is intended for integration of monoscopic sequences into stereoscopic presentations. The same monoscopic image is used instead separate left and right images to calculate the stereoscopic output. Stereoscopic Player allows adjusting the horizontal parallax to move the scene behind or in font of the screen as well as adjusting vertical parallax to compensate for vertical misalignment. All adjustment features are available for monoscopic content, too. Parallax adjustment is the most useful one to create a pseudo-3D effect. Furthermore, Stereoscopic Player can flip left and right images horizontally and vertically. Supported output formats include:

– Monoscopic
– Dual screen output
– StereoBright dual screen output[*]
– NVIDIA stereo driver
– Quad-buffered OpenGL
– Side-by-side
– Over/under
– Row interlaced
– Column interlaced
– Pure, grey and color anaglyphs for red-blue, red-green, red-cyan and yellow-blue glasses
– Half-color anaglyphs for reduced retinal rivalry
– Novel retinal rivalry free optimized anaglyphs

The NVIDIA stereo driver viewing method is the most flexible method supported by Stereoscopic Player. The stereo driver is available for all NVIDIA graphics cards, including the consumer GeForce series; it must not be confused with quad-buffered OpenGL stereo support. When the NVIDIA stereo driver viewing method is selected, Stereoscopic Player passes unaltered left and right image buffers to the driver. Depending on the driver's configuration, pageflipped output or color anaglyph output will be created. Undocumented driver features provide additional methods, for example dual screen output, horizontal and vertical interlaced output and a proprietary format for Sharp's autostereoscopic LCD displays. The interlaced output methods can be used instead of Stereoscopic Player's own viewing method to achieve better performance. Whereas NVIDIA's stereo driver works in fullscreen mode only, quad-buffered OpenGL provides in-window pageflipping. Similar to NVIDIA's consumer stereo driver, OpenGL drivers may provide additional output formats beyond pageflipping. However, quad-buffered OpenGL support is available for professional graphics cards only, for example for NVIDIA's Quadro FX series.

---

[*] StereoBright is a passive polarized projection method for LCD projectors using half-wavelength retardation films. Contact Advisol for more details (http://www.advisol.co.il).

3. USABILITY ISSUES

Stereoscopic video playback burdens the user a lot of configuration tasks compared to monoscopic playback. In addition to problems related with video playback in general (missing codecs, performance issues, ...) the proper layout must be chosen, where it is most difficult to distinguish similar layouts, e.g. side-by-side left/right and side-by-side right/left. Especially people new to stereoscopy fail to recognize pseudostereoscopic images and to choose the reverse image order, resulting in a bad first time experience. The separation between the left and right images and the aspect ratio are other options which must be set before starting playback.

Stereoscopic Player implements several mechanisms to simplify these tasks. In the first place, it remembers settings for each video in its built-in library (see Figure 2). In addition, Stereoscopic Player supports configuration files, called stereoscopic video information files, which contain all information necessary to play the matched video file. The stereoscopic video information file (file extension: svi) and the video file must be located in the same directory and share the same filename. Stereoscopic Player automatically detects svi-files when loading the associated video files. Downloading stereoscopic information files from a web server is supported as well. This Internet service is entirely based on the hypertext transfer protocol (HTTP). Stereoscopic Player performs the following steps to retrieve information from such a web server:

1. Stereoscopic Player requests UpdateLibrary.dat and validates its header to verify that the specified web address actually provides stereoscopic information files.

2. Stereoscopic Player calculates a 64 bit hash id based on the content of a video file or a DVD and uses this number to request the stereoscopic information file. The hash value in decimal format is directly used as the filename. The detailed description of the hash calculation algorithm can be found in [8].

3. If the server returns an error during step 1 or 2, the algorithm continues at step 1 trying the next server, if available. By default, Stereoscopic Player is configured to use the official web server only, but additional servers can be added easily.

Stereoscopic information files and Stereoscopic Player's Internet service transfer configuration workload from the users to the authors of stereoscopic videos. This makes stereoscopic video playback almost as easy as monoscopic video playback. The official web server already provides settings for over hundred stereoscopic DVDs and video files.
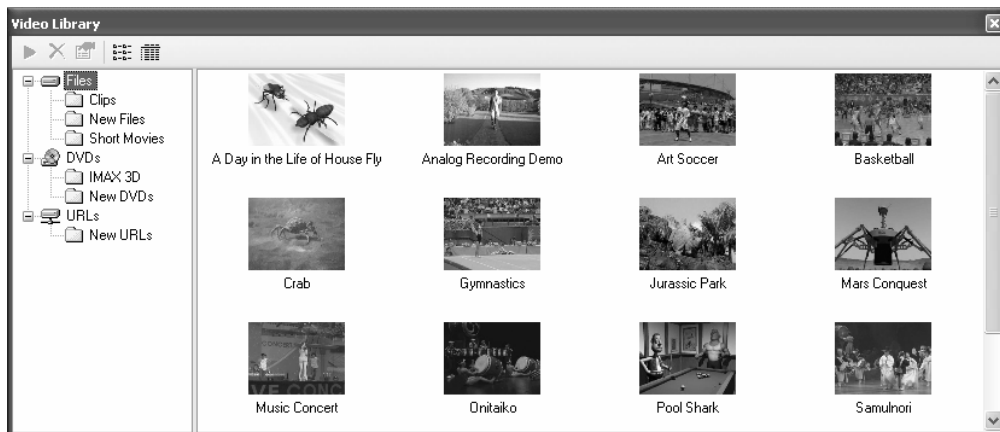

Figure 2. Stereoscopic Player's video library

4. PLAYER ARCHITECTURE

Stereoscopic Player is based on Microsoft's DirectShow, a modular architecture to handle streaming data. Processing is done by objects, called DirectShow filters, which are connected to a directed graph controlled by the filter graph manager. DirectShow provides a set of standard filters, including source filters, decoders and encoders for most popular video formats. It also provides audio and video rendering filters as well as access to capture hardware. DirectShow is a part of DirectX. The DirectX SDK[5] allows developers to write their own filters in order to extend DirectShow's functionality. Filters expose one or more input or output pins to connect with other filters. During the connection process, filters must agree on a common media type.
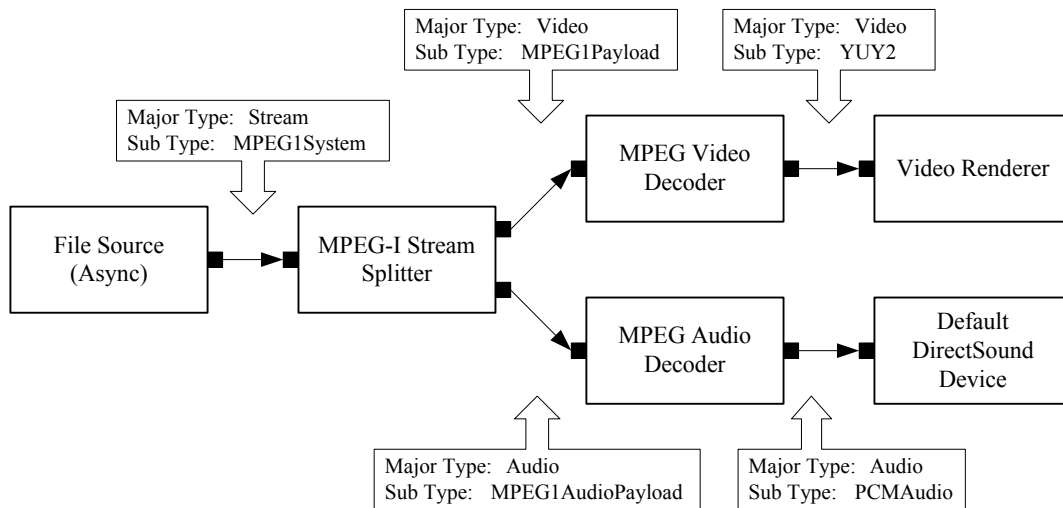

Figure 3. MPEG file playback graph

Figure 3 shows a graph to play an MPEG file, including annotated connection media types. This graph must be modified for stereoscopic video playback, since we have to transform the video data depending on the desired viewing method. To do any kind of image processing, we need uncompressed image data. Therefore, the only reasonable position to modify the graph is downstream the video decoder. For most stereoscopic viewing methods, inserting a transformation filter between decoder and video renderer is sufficient (see Figure 4). Only quad-buffered OpenGL and NVIDIA stereo driver support require special renderers, replacing DirectShow's default video renderer. Consequently, Stereoscopic Player uses three own filters to implement stereoscopic playback: the *Stereo Transformation Filter*, the *OpenGL Stereo Renderer* and the *NVIDIA Stereo Renderer*. Microsoft provides a helpful tool called GraphEdit to manually build and test filter graphs. It allows using these three filters to do special stereoscopic processing tasks not covered by the player user interface.

The Stereo Transformation Filter processes all frames at their native resolution. For example, if the video is encoded in over/under format at 640 x 480 pixel (640 x 240 pixels per eye), anaglyph output resolution would be 640 x 240 pixels and side-by-side output resolution 1280 x 240 pixels. Upscaling and stretching to obtain the desired resolution and aspect ratio is done by the video renderer using DirectX hardware acceleration. Only interlaced output must be created at the final resolution since any stretching operation later on would destroy the interlaced pattern.

DirectShow supports numerous uncompressed video data formats. Among the most important ones are RGB 24, RGB 32, YUY2 and UYVY, all of them supported on Stereo Transformation Filter's input pin. RGB 24 and RGB 32 are both 24 bits per pixel RGB formats, aligned on 24 bit and 32 bit, respectively. YUY2 and UYVY are 16 bits per pixel, 4:2:2 color subsampled YUV formats, differing in byte order only.
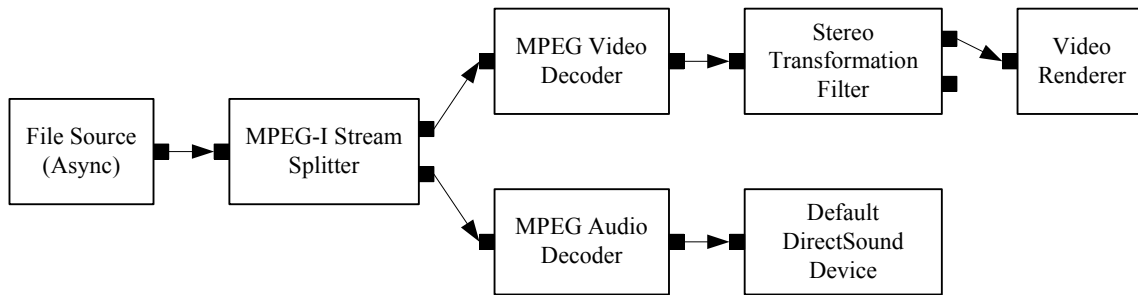
Figure 4. MPEG file playback graph, modified for stereo stereoscopic output

The second output pin of the Stereo Transformation Filter is required for dual screen output. When playing the video in side-by-side format and maximizing the playback window over two screens, one screen stays black on most graphics cards because hardware overlays are restricted to only one screen. To work around such problems, Stereoscopic Player provides the dual screen output method which utilizes two separate video renderers for left and right screens.

Figure 5 shows a 3D-DVD playback graph. Whereas the DVD Navigator is provided by DirectShow, third party video and audio decoders are required. Stereoscopic Player's configuration dialog allows selecting any DirectShow-compatible decoders. Free suitable video and audio decoders are the GPL MPEG Decoder[6] and AC3Filter[7], respectively. Internet streaming is implemented similarly to file playback. Instead of a file source, a network source filter must be used. By default, DirectShow offers a source filter to receive Windows Media streams. Any other format can be added without modifying Stereoscopic Player by installing appropriate third party source filters. Capture devices are the fourth supported video source. Both WDM and legacy Video for Windows drivers can be used and all driver configuration dialogs are accessible via the player's file menu.
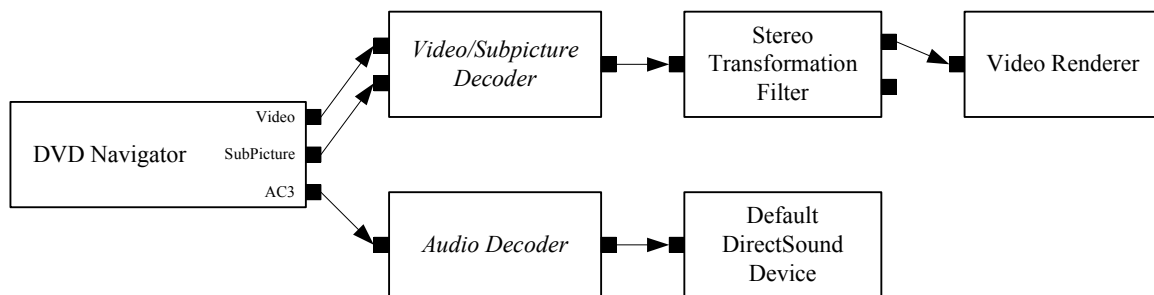

Figure 5: Stereoscopic DVD playback graph

5. STEREOSCOPIC VIDEO RECORDING

There are two main options in recording high quality stereoscopic videos: (1) using a specialized camera or (2) using two standard cameras. While the second option is expected to have significantly lower costs for the hardware equipment, it is quite challenging with respect to synchronization and adjustment of the two cameras. This paper discusses the second option only.

We distinguish two scenarios of operation: In the first case, camcorders record the left and right video streams to their built-in storage media, e.g. tapes. The final stereoscopic video is assembled later on during the post production process. The second scenario—live transmission or computer-based hard disk recording—requires both streams to be captured, assembled and (re-)compressed immediately.

Our stereoscopic recording solution was designed to fit the needs of both scenarios. Stereoscopic Multiplexer is a driver that can be used within many applications to capture stereoscopic video content. It behaves like a WDM driver, and is therefore compatible with applications supporting WDM drivers. Stereoscopic Multiplexer acquires video frames from two 'real' capture devices, combines them to side-by-side format and passes the resulting sequence of stereoscopic images to the application. The novel implementation allows the multiplexer to run entirely in the operating system's user mode. Execution in user mode, in opposite to kernel mode, cannot compromise system stability.

Investigations of several video applications designed for Windows XP proved that none of these applications directly interfaces the driver. Instead, they rely on DirectShow as an intermediate layer to access capture devices. Consequently, Stereoscopic Multiplexer is implemented as a DirectShow filter, not as a WDM driver. It exposes all interfaces provided by the WDM Capture Filter, which encapsulates WDM drivers, and is registered as a capture device. As a result, applications using DirectShow to enumerate and access capture devices will see the Stereoscopic Multiplexer and cannot distinguish this virtual capture device from a real piece of hardware.

Whenever an application creates a capture graph, it loads the capture driver and all necessary downstream filters. However, when using the Stereoscopic Multiplexer, it will never load the real capture devices. It is the multiplexer's responsibility to reload the left and right capture drivers when necessary. This is done when its output pin gets connected. Figure 6 shows a typical stereoscopic capture graph. The 'Sink' box is a placeholder for several other downstream filters, e.g. video renderers, A/V multiplexers, file writers and many more.
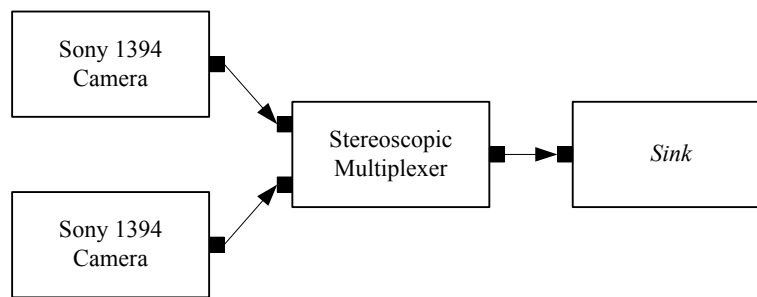

Figure 6. Example for a stereoscopic capture graph

Three software layers—illustrated in Figure 7—are involved in video capturing: the application layer, the DirectShow layer and the driver layer. When comparing monoscopic and stereoscopic capturing, shown on the left and right side of Figure 7, respectively, we see that the Stereoscopic Multiplexer introduces an additional layer between the application and the capture drivers. From the application's point of view Stereoscopic Multiplexer acts as a capture filter; from the driver's point of view it replaces the application layer.

Stereoscopic Multiplexer currently supports RGB 24, RGB32, YUY2 and UYVY formats. In both YUV modes, it can apply gamma correction, unsharp masking, white and black level adjustments and conversion to ITU-R BT.601 colorspace[*] in real-time. In addition, it can flip left and right images both horizontally and vertically, which is useful when mounting one camera upside down to minimize stereo base, the distance between left and right camera. Ideally, the capture drivers should deliver uncompressed data, so that no decompression to any of the supported formats must take place before multiplexing. Saved CPU power can be used for further processing and efficient compression necessary for Internet broadcasts.

---

[*] ITU-R BT.601 colorspace is adopted by all major DirectShow YUV formats. Luminance values are in the range between 16 and 235, chrominance values between 16 and 240. However, some cameras use the full dynamic range from 0 to 255. Without conversion, contrast would be exaggerated and details in bright and dark areas would get lost.
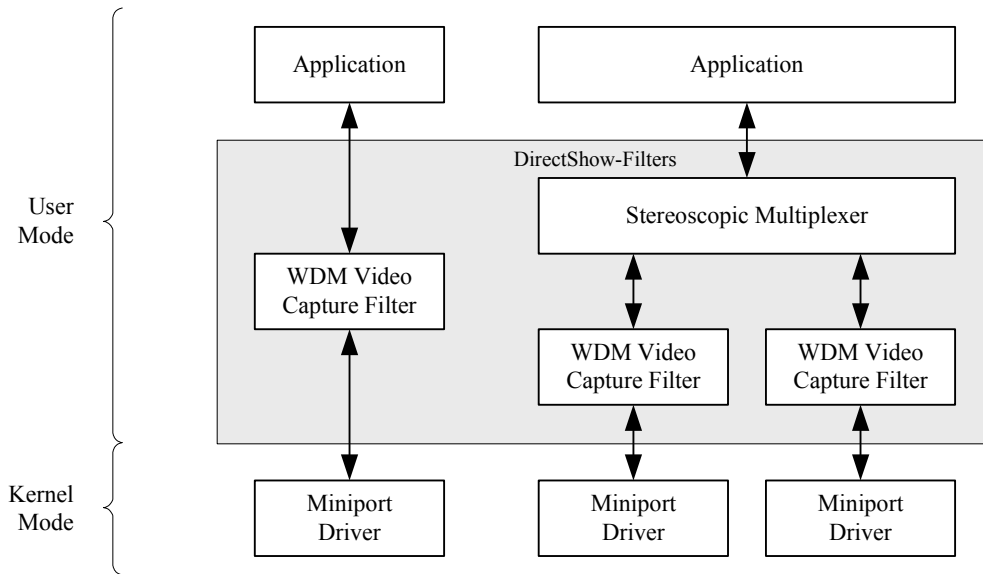
Figure 7. Software layers involved in monoscopic and stereoscopic capturing

6. SYNCHRONIZATION ISSUES

Each media sample (for video data, a media sample corresponds to one frame) traversing the filter graph contains a timestamp specifying its start and end time. Based on the timestamp information, Stereoscopic Multiplexer combines left and right frames with minimal time offset. Optimal synchronization requires the left and right sources to have the same framerates. In this case, mis-sync is lower or equal to half the duration of a frame. Obviously, software synchronization cannot provide zero mis-sync—this would require hardware genlocking. When the framerates of left and right camera differ, frames are dropped periodically to obtain stereo pairs matched as well as possible. Variable framerates especially occur with USB 1.1 webcams due to limited USB bandwidth. USB 2 and Firewire cameras usually deliver frames at a stable framerate.

Synchronization based on DirectShow's timestamps does not take into account that timestamps need not be correlated with the actual recording time. Timestamps are assigned by the capture driver and do not include any latency in the camera or on the bus. Our tests showed that timestamps provided by the Microsoft DV capture driver are virtually useless for synchronization (no matter which DV camera is actually used, even when the cameras are synchronized in hardware), whereas for example two Sony DFW-VL500 industrial Firewire cameras provide accurately matched timestamps.

Several webcams and industrial Firewire cameras allow adjusting camera parameters (brightness, contrast, color balance, exposure, zoom, focus, ...) in the driver configuration dialog. Usually, these settings can be changed programmatically through standardized DirectShow interfaces as well. Stereoscopic Multiplexer relies on these interfaces to provide a user interface to configure both cameras simultaneously (see Figure 8). Some camera parameters support an automatic mode. In such cases, Stereoscopic Multiplexer activates the automatic mode only on the first camera, periodically reads out the value and transfers it to the second camera. Manual parameters will be synchronized in the opposite direction as well. Therefore, manual changes on one camera are periodically transferred to the other camera. Unfortunately, some cameras don't allow controlling their parameters numerically and consequently do not provide certain DirectShow-interfaces required by Stereoscopic Multiplexer's parameter control feature. Among these devices are the well known DV cameras.
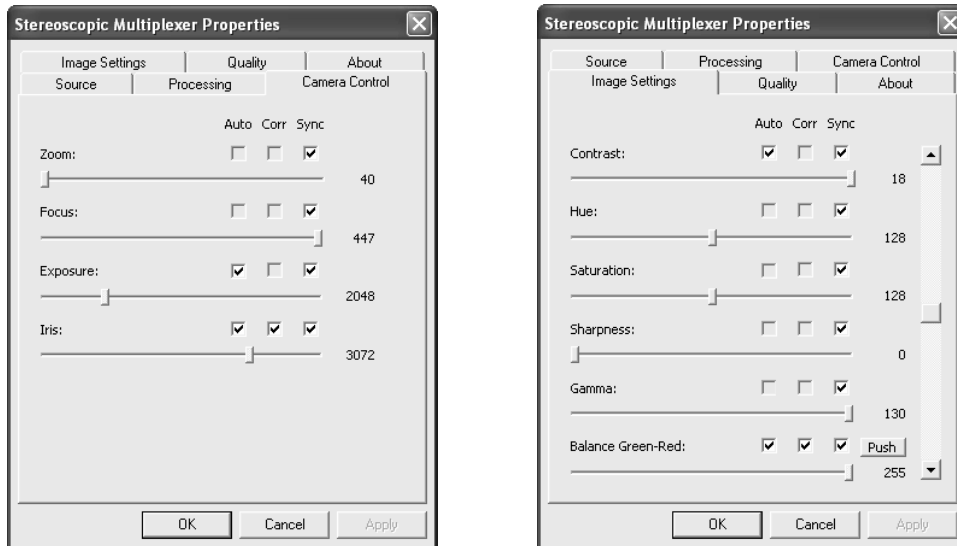
Figure 8. Stereoscopic Multiplexer property pages (showing parameters supported by Sony DFW-VL500 cameras)
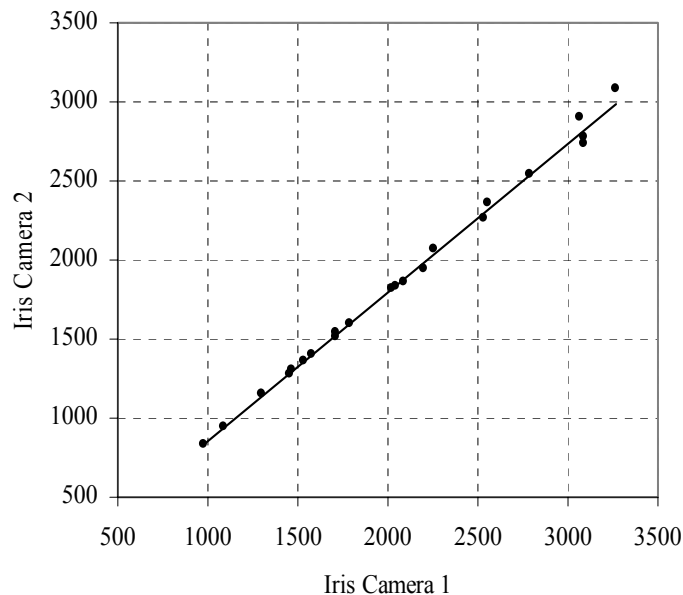

Figure 9. Relationship between left and right iris parameters to obtain equally bright images

Despite same settings on both cameras, significant discrepancies between left and right image might occur. They are caused by manufacturing tolerances and different factory-made adjustments. In our recordings, shot with Sony DFW-VL500[*] cameras, different brightness and different color balance were quite obvious. An analysis showed that it is possible to eliminate all discrepancies by varying iris, green-red balance and yellow-blue balance settings. We corrected

---

[*] The Sony DFW-VL500 is a progressive scan Firewire camera supporting a maximum resolution of 640 x 480 at 30 frames per second. It delivers image data in uncompressed YUV 4:2:2 or 4:1:1 format. We decided to use DFW-VL500 cameras because the built in 12x zoom lens can be set to any zoom and focus values in software and all other parameters can be computer-controller as well.

iris and color balance settings manually under different lighting situations to retrieve a correction table for each parameter. Figure 9 shows the relationship between iris settings for left and right camera to get equally bright images.

Stereoscopic Multiplexer has been extended to use correction functions whenever parameters are set or synchronized. Although a single linear function would have been sufficient to perfectly synchronize our cameras, we decided to implement a more generic parameter synchronization algorithm. A piecewise linear, continuously increasing[*] function can be specified for each parameter. Whenever a parameter on the right camera is set, the correction function is applied first. In case a parameter on the right camera has changed, it is copied to left camera by applying the inverse function.

7. MEASURING MIS-SYNC

As mentioned before, we cannot rely on timestamps to find out the mis-sync of multiplexed stereo pairs. Instead, we have to record a well known motion which allows us to calculate the time offset based on different stages of motion. Because real moving objects are difficult to handle, we decided to use a CRT monitor's electron beam as reference motion. It draws the image line by line starting at the top left corner. For this purpose, the test image seen in Figure 10 is displayed on the screen. The scale allows reading off the vertical position in pixels. Additional vertical bars are added after each refresh cycle so that we can distinguish consecutive refresh cycles later on. Displaying the vertical bars cannot be done by playing a pre-created video file but requires a special application that guarantees that no cycle is missed.
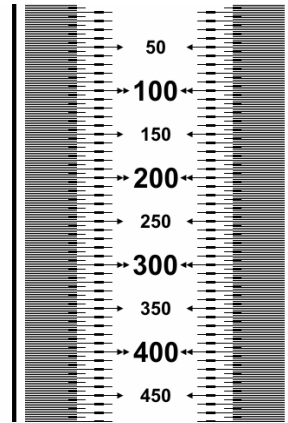

Figure 10. Test image used for mis-sync measurements

The mis-sync can be measured by filming the screen using a very short exposure time. Figure 11 shows a stereo pair recorded with Sony DFW-VL500 cameras at an exposure time of 1/256 seconds. As we can see, the right image was recorded slightly earlier, since the electron beam has just reached line $z_1 = 223$ whereas it is already at line $z_2 = 403$ on the left image. Both images where taken during the same refresh cycle because the number of vertical bars equals.

To calculate the mis-sync t_{Offset}, we need to know the current VGA timing. Most monitors display the horizontal refresh rate f_{Horiz} and vertical refresh rate f_{Vert} in their on-screen displays. d denotes the difference of vertical bars between left and right image.

$$t_{Offset} = \frac{z_2 - z_1}{f_{Horiz}} + \frac{d}{f_{Vert}} \tag{1}$$

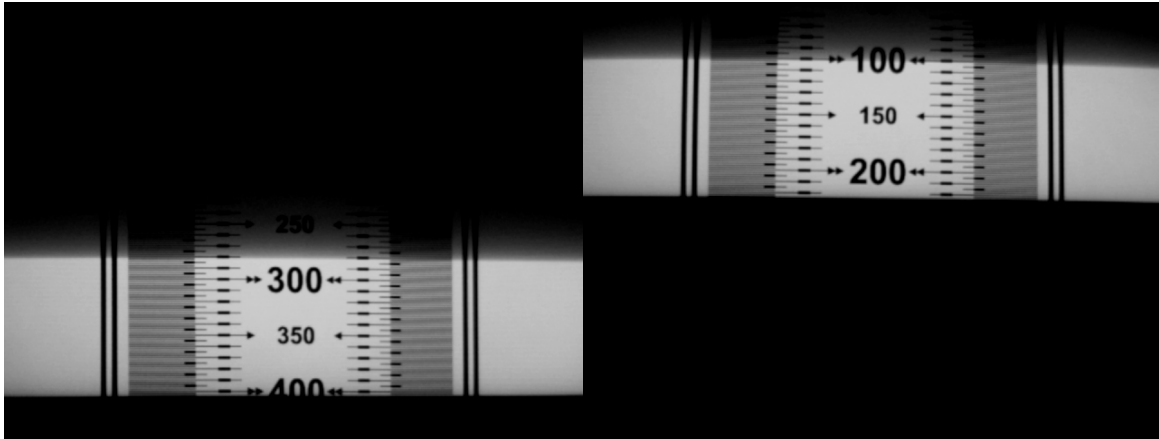[*] Necessary so that the inverse function exists.

Figure 11. Resulting stereo pair after filming the test image displayed on a CRT monitor

In our example, the screen was set to the standard VGA mode at 60 Hz (f_{Horiz} = 31.5 kHz, f_{Vert} = 60 Hz) when taking the images shown in Figure 11. Putting in these values, formula (1) evaluates to:

$$t_{Offset} = \frac{403 - 223}{31500} + \frac{0}{60} = 5.7\,\text{ms} \tag{2}$$

We did a series of tests with a pair of Sony DFW-VL500 cameras (see Figure 12). It turned out that synchronization based on timestamps works well with these cameras. Mis-sync of recorded clips was randomly distributed in the range between 0 and 1/60 second in all test recordings. Finally, we tried to further improve synchronization by powering on both cameras simultaneously. However, instead of perfect synchronization, we got an average offset of 10 ms (average deviation 0.25 ms). 10 ms offset is already problematic when recording fast moving objects, but it is acceptable for low motions scenes, for example video conferencing applications.

Figure 12. Pair of Sony DFW-VL500 cameras

8. CONCLUSIONS

Stereoscopic Player is a universal solution to render stereoscopic videos. Due to its variety of supported formats and media sources (files, DVDs, web streams, capture devices), its easy to use interface and its rich features, Stereoscopic Player is prepared for many scopes of application. It fulfills the needs of both consumers and professionals. Since

Stereoscopic Player uses a modular concept based on DirectShow, it can be easily extended to support new codecs and stereoscopic viewing methods. The second solution presented in this paper—Stereoscopic Multiplexer—combines two live video streams to a single stream in side-by-side format while minimizing time offsets. Furthermore, it synchronizes camera parameters and offers image processing features. Typical fields of application include:

1. Stereoscopic real-time preview during recording
2. Stereoscopic hard disk recording
3. Stereoscopic video transmission over the Internet

Suitable cameras for application (1) would be DV camcorders, whereas webcams or industrial cameras delivering uncompressed data are best suited for applications (2) and (3).

ACKNOWLEDGMENTS

REFERENCES

1. Stereoscopic Player website: http://mitglied.lycos.de/stereo3d
2. Stereoscopic Multiplexer website: http://mitglied.lycos.de/stereo3d/multiplexer.htm
3. L. Lipton, Foundations of the Stereoscopic Cinema, Van Nostrand-Reinhold, New York, 1982
4. M. Hušák, C. Ward, "DepthQ: universal system for stereoscopic video visualization on WIN32 platform", Stereoscopic Displays and Virtual Reality Systems XI, Proc. of SPIE Vol. 5291A, pp. 314-320, 2004
5. Microsoft DirectX SDK website: Website: http://msdn.microsoft.com/library/en-us/dnanchor/html/anch_directx.asp
6. GPL MPEG Decoder website: http://sourceforge.net/projects/gplmpgdec
7. AC3Filter website: http://ac3filter.sourceforge.net
8. P. Wimmer, "Aufnahme und Wiedergabe stereoskopischer Videos im Anwendungsbereich der Telekooperation", master thesis, 2004