Avid's Stereoscopic Editorial Architecture

Light Fields, Intelligent Computing and Beyond

Shailendra Mathur, Video Chief Architect | March 19th 2013
Topics

- Stereo 3D data handling
- Avid Intelligent Compute Architecture
Avid introduction

- Tools for Film, Broadcast and Live productions
- Video and Audio
- Focus today on Film editorial editing tool – Media Composer
- Our customers are Academy, Grammies and Emmy award winners, • ... and so are we!
Stereo 3D Data Handling
Trends

- Increasing granularity (resolution) of video
  - High frame rate usage
  - High spatial resolution
  - High bit-depth
- Increasing number of views
  - Stereo shooting - Mono and S3D deliverables
  - Plenoptic/Light Field Cameras
  - Camera Grids
“Super-sample the work, edit it later!”

- The Story-teller editing the film wins!

- More elements of the scene to work with
  - Zoom-in,
  - Pan and scan
  - Slow-down,
  - Speed-up
  - Increase contrast,
  - Expose new dynamic ranges

...without losing display precision!
Problems to solve

- Need a data model for managing multiple views and resolutions
- Support common editing and data management functions and workflows
Light Fields – a parametric model

- 7-D Plenoptic Function
  \[ I = P(\Theta_v, \Phi_v, E_x, E_y, E_z, t, \lambda) \]

- Camera streams are spatio-temporal-colorimetric view-samples of

  “Radiance received along any direction V arriving at any point E in space, at any time ‘t’ and over any range of wavelength”
Multi-view sampling of a light field

http://commons.wikimedia.org/wiki/Main_Page
Multi-cam productions

Temporally aligned multi-cam group

Time Reference
Registered multi-cams

Scene

- Color synchronization
- Spatial synchronization
- Temporal synchronization
Multi-channel video data-model
Proxy Editing and Multi-Mastering

Transcode to side by side S3D

Capture full frame uncompressed S3D left and right views

Output context: SideBySide, L|R, DNX36, L|R audio

C1: Left Eye (Leading)  
C2: Right Eye

Channel Combiner

Output context: Full Frame, L|R, uncompressed, L|R audio

Stage 1

Stage 2

Stage 3

Full Frame S3D Master output

Mono Master output using leading eye

Output context: Full Frame, Leading Eye, DNX444, L|R audio

S4 with Right Audio

S5 with Right Audio

S4 with Left Audio

S3 R|L 1:1

Resolution 2

S2 R 1:1

Resolution 1

S1 L 1:1

Resolution 3

S4 L DNX444
Data Model – Problem solved.

- Method of grouping multiple views
- Spatial, temporal and Color synchronization of views
- Method of arranging resolutions under the views
- Resampling of scene based on output context.
Avid Intelligent Compute Architecture
Performance Issues

- Multiple views => more processing
  - codecs,
  - visual effects
  - graphics,
  - visualization

- Higher resolution, frame rates and bit-depth
Avid Intelligent Compute Architecture

- Heterogeneous compute Hardware
  - Intel multi-core CPUs (HP Z800/820, Apple Mac Pro)
  - Nitris DX w/ dual DNxHD or AVC-’ codec acceleration modules
  - NVIDIA GPU
  - NVIDIA 3DVision support

- Intelligent Media Player
Problem with heterogeneous compute

- Multiple pieces of code for each compute domain
- Balancing across multiple compute domains dynamically
- Optimal transfers between hardware?
- Dynamically adding a new hardware processing domains
How others have solved it

- **Main philosophy - "Write once, support many"**
  - Generic Language with HW specific compilation, e.g. OCL
  - Code generation for specific compute models, e.g. Cg
How we solved it

- Philosophy - "Optimize Many, Host once"
  - write compute domain specific optimized code
  - single API contract with client application.
  - Domain-abstracted player
Plug’N Play Architecture

- Home grown component model
  - Similar to COM

- Plug and play
  - Compute domains
  - Processors
  - Data types and formats
  - Converters
Plugging-in NVIDIA GPU - Benchmark results

- System
  - Avid Symphony 6.0
  - HP Z800 Workstation
  - Windows 7 64-bit
  - 6 GB RAM
  - NVIDIA Quadro 4000
  - Avid Nitris DX
  - Avid VideoRaid storage

- Material
  - AVC-Intra 100 sources rendered to Avid DNxHD 185x
  - 1080i/50 10-bit
  - Processing:
    - Full resolution HD
    - Full-Frame processing of Stereo 3D streams 16 bit processing precision
    - Render quality: best (advanced polyphase)

<table>
<thead>
<tr>
<th>Task</th>
<th>How much faster with NVIDIA GPU?</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 streams with title</td>
<td>2.6x</td>
</tr>
<tr>
<td>4 color corrections, 3 picture-in-picture with soft border, 1 title</td>
<td>2.6x</td>
</tr>
<tr>
<td>4 streams with title</td>
<td>7.3x</td>
</tr>
<tr>
<td>4 color corrections, 3 perspective (3D Warp) with soft border, 1 title</td>
<td>7.3x</td>
</tr>
<tr>
<td>2 advanced effect streams</td>
<td>1.8x</td>
</tr>
<tr>
<td>1 color correction + stabiliser + spectramatte key + resize + Symphony advanced color correction</td>
<td>1.8x</td>
</tr>
<tr>
<td>1 FullFrame stereo effect stream</td>
<td>1.4x</td>
</tr>
<tr>
<td>color alignment + positional alignment + S3D rotational alignment + vergence + zoom + floating window</td>
<td>1.4x</td>
</tr>
<tr>
<td>2 FullFrame stereo effect streams</td>
<td>1.9x</td>
</tr>
<tr>
<td>2 color alignments + 2 positional alignments + 2 S3D rotational alignments + Depth transition + Dissolve</td>
<td>1.9x</td>
</tr>
</tbody>
</table>
Example
AVC-I Decode + color alignments + 2 positional alignments + 2 S3D rotational alignments + Depth transition + Dissolve + DN220x encode

- Stereo Full Frame AVC-I decode
- color alignments per stereo pair
- positional alignments per stereo pair
- S3D rotational alignments per stereo pair
- Depth transition
- Dissolve
- DNx220x 10bit Encode