

VFX By the Numbers: A Technical Introduction to Previzion

Eliot Mack
CEO
Lightcraft Technology
June 20, 2011

Origins

Like all good projects, the development of Previzion began with personal frustration at the status quo; I helped a friend work on his student film just as the consumer DV revolution started. The ease of digital capture made the future of the industry obvious, but there was a clear gulf between the look of the images one could capture on a miniDV recorder, and what a DVD of a 'real' movie looked like, even though the pixel counts were identical between the camcorder capture and the DVD. I wanted to know why, and finding the answers to that little question turned into an eight year long project.

Small budgets, big ideas; this was not a new desire. Some parts of creating production value were obvious; set extensions with visual effects could turn a small set into a large space. I taught myself the basics of VFX creation pretty quickly, and completed all of about two shots before becoming exasperated with the state of the technical toolset. Tasks that I thought would be simple (pulling a green screen key, tracking the movement of a camera) turned out to be very difficult. At my current rate of execution speed, I realized, it would take hundreds of people years to make a movie, which turned out to be quite true.

By that point, however, I understood part of the math underlying visual effects, and thought that I could build a system that would take care of the boring, repetitive tasks for me. I had just finished engineering the machine design of the world's first successful home robot, the iRobot Roomba, and thought I could solve most of the fundamental problems in visual effects in a couple of years. This turned out to be more than a little naive, but in retrospect it was better not to know what I was getting into.

I discovered that real time visual effects involved several very hard problems at the limits of image processing technology, including keying, photorealistic 3D rendering, inertial navigation, and lens calibration. All of these technologies had to operate on real sound stages and on locations, which are about as far from laboratory conditions as it is possible to get. Since the motion picture and television production markets are so small, very few off the shelf technologies had been developed that could handle their problems, so everything had to be built specifically for the market.

After the initial release of Previzion, I was further surprised to see that there were two distinct groups that began to use the system. As previzualization for feature films and high end episodic television production was an established use for this type of system, the needs of accurate preview and data recording were expected. What surprised us was the demand for real time, in camera finishing. For many production groups, such as telenovelas and soap operas, the rate of production speed is so high that traditional VFX pipelines just can't keep up. Based on requests from these groups, the Previzion toolset evolved to handle both of these areas of work, with the underlying connecting thread of high accuracy and image quality.

Subsystems

At its core, Previzion is a visual effects pipeline that runs in real time. It can be used either as a visual effects preview tool that records camera tracking data for use later in post production, or as a real time finishing tool for in camera compositing. It is based on programmable hardware from the input video connector to the output video connector, which enables feature upgrades by simply adding more lines of code. Previzion uses a combination of parallel CPU, GPU, and FPGA processing to handle the various types of computation challenges posed by the needs of real time processing. This renders the system reasonably future proof; the underlying hardware can be upgraded every couple of years, while the software can be upgraded continuously.

Previzion is designed with multiple stages that closely match each part of a traditional visual effects pipeline. These include real time camera and lens tracking, keying, 3D rendering, compositing, live and pre-recorded video I/O, animation I/O, and data recording. The various data formats are generally based on open XML formats such as Collada and standard interchange formats such as FBX, which makes it straightforward to integrate Previzion with existing real time and post production systems.

Camera and Lens Tracking

The camera tracking stage works with a wide variety of tracking systems, including the inertial/optical Airtrack/Intersense combo, encoded Technocranes and sliders, GPS/IMU outdoor navigation systems, and motion capture stages. Nearly any real time 6DOF data tracking source can be added quite quickly, and a coordinate transformation toolset lets users register the tracking data into a local surveyed coordinate space very quickly, so that CG objects line up correctly with live action objects.

We put considerable work into solving the camera tracking problem for everyday tracking on typical sound stages, as opposed to needing special motion capture stages. This involved building our own high accuracy rotary gyro, the Airtrack, with hardware synchronization to video genlock sources. This provides data in perfect sync with the camera motion, which is critical to keep the virtual background moving seamlessly with the foreground.. In this case, we combine the high accuracy rotary data from the Airtrack with the position data from the Intersense IS1200 smart camera to provide an overall 6DOF solution that works with simple printed paper targets mounted on the ceiling of soundstages. This tracking system can be rigged on a stage and calibrated in a single tech setup day, and works without interference from existing stage lighting methods.

We also recently added GPS/IMU tracking integration. This type of tracking integrates GPS satellite tracking units, a six degree of freedom IMU/accelerometer, and a local differential GPS station to provide centimeter-level tracking accuracy in exterior locations, with no need for tracking markers or encoded cranes. Since the tracking can work anywhere within a kilometer of the base station, this has enormous potential for exterior on location shoots incorporating a lot of VFX.

The lens calibration also required inventing a whole new technology to handle the unique demands of VFX production. To correctly match CG and live action together, the system must completely map the lens parameters, including the true focal length, the focus distance, the amount of lens distortion, and the entry pupil location. This calibration process needs to cover the wide range of focal lengths and

parameters found in modern cinema zoom lenses. Existing lens calibration processes for virtual studios typically require a dedicated calibration room with non-portable calibration targets, and a separate software tool.

Previzion's unique, patent pending lens calibration system uses a moving, motorized lens calibration target that moves at varying distances from the camera lens to resolve the various lens calibration parameters. The calibration target is flat, and is small enough to fit in a standard van or station wagon. This lets the lens calibration process happen on the stage, wherever the lens and camera happen to be, and transforms a process that typically required shipping the lens out to a calibration facility for a couple of weeks into a process that is done in 3 hours during a normal stage tech preparation day. The lens calibration file is saved after completing calibration, and can be reloaded instantly for daily usage.

The lens calibration system also uses standard production lens controllers to measure the lens barrel motion, such as encoded Fujinon and Canon lenses, as well as Preston and RT Motion lens controllers. This way, the lens data is read using systems that production crew are already familiar with.

Video Processing

Previzion's real time video system handles both multiple live and pre-recorded video sources and outputs. The live video system has 2 HDSDI inputs and 2 HDSDI outputs. One video input drives the keyer, and the other can be mapped into any texture in the 3D scene designated as a video texture. The video outputs can show the untouched green screen foreground, the background rendered scene, the complete composited image, or a matte signal from the keyer.. All of the outputs have SMPTE timecode embedded in the HDSDI VANC, so that the signal can be recorded and used immediately by editorial and compositing with a frame-accurate matchback to the original greenscreen signal. Timecode can be read either directly from data embedded in the HDSDI VANC signal, or read from external LTC analog timecode sources.

In addition, Previzion is designed with a parallel processing video decoder that can decode up to 8 streams of 1920x1080 HD MPEG-2 video that is then mapped to selected textures in the 3D scene. In this way, live action background plates can easily be combined with 3D midground and foreground elements as per traditional VFX techniques.

Previzion can even operate in interlaced rendering mode with interlaced video inputs and source footage, where the scene is rendered at 59.94 frames/second using odd or even lines from the MPEG footage to make the background plates match perfectly with the foreground camera motion and preserve the look of interlaced footage.

As noted above, Previzion's 2D and 3D subsystems are linked together and programmable. This means that the live action can be located behind foreground CG objects, such as putting the driver in a vehicle behind a 3D virtual windshield with transparency and reflections rendered on the windshield.

3D Rendering

Previzion includes a real time 3D rendering engine that is synchronized to the camera motion. The renderer can read user-created CgFx real time surface shaders for precise description of various materials, such as reflections, bump, and normal maps, and even stitch together multiple video plates in real time. The renderer also includes photographically accurate depth of field and motion blur, which takes into account the focal length, focus distance, sensor size, and current F stop so that the cinematographer can rack focus back and forth between the CG objects and the live action objects just as if they were shooting on a real scene.

The 3D virtual scene is loaded with a custom FBX or Collada file loader, which lets users export 3D sets out of standard content creation tools such as Maya and 3ds max, including high resolution textures, UV maps, embedded surface shaders, and lighting. This drastically simplifies the content creation pipeline, as the tools that are used by 3D artists to create content for Previzion are the tools that they are already accustomed to using.

In addition, lens distortion is both measured and used when rendering the 3D virtual background; this is required if objects in the virtual space are to be matched with objects in the live action space.

Keying, Compositing, and Color Correction

Previzion uses a VFX quality keying algorithm adapted to run in real time, with additional custom controls to handle the wide range of keying situations found in real world productions. The dvMatte Pro algorithm combines a powerful color difference keyer along with a matteless compositing stage that can extract single flyaway hairs and preserve them in a composite, without color spill.

Multiple edge and matte adjustment controls, including matte blurs, lateral and vertical shifting, edge coloration, and add/mix composite controls allow the operator to handle the majority of tricky edge cases that occur in production.

In addition, separable foreground and background color correction stages let the user bend the foreground and background coloration to match the white points, the black points, and the color levels, to provide a seamless photographic match.

Live Animation

Previzion's open architecture also extends to its integration with external animation systems. Real time streams of animation data can be input or output from other animation tools. This lets users connect live animation sources, such as motion capture suits, and use them to drive animated characters in the scene. In addition, Previzion generates a real time output data stream of the current frame's camera data, including the current timecode value and the complete camera position, orientation, and lens parameters needed to recreate the rendered scenes. This lets the system connect easily to external engines. The TCP/IP network animation protocols corresponds closely to the Collada XML standard format, so that external tools can easily be made to interface with Previzion.

Metadata Logging and Data Transfer

Previzion contains a very complete data logging system. The user simply presses the Roll button to start recording, and Cut to stop recording. Each take is recorded into a Collada XML text file, which is both human readable and machine readable. This file contains the complete camera position, orientation, and lens data for each frame, along with the matched timecode and shot details like frame rate, sensor size, and the current timecode of incoming animation data from Motionbuilder sourced animations.

Some software packages can read the Collada format directly. However, since many 3D post production packages have their own preferred software format, Lightcraft built a program called Take Converter that reads standard Previzion generated Collada take files, accepts an in and out timecode to specify which part of the take data is required, and exports just those selected frames to whatever format is requested by the user (Maya ASCII, Nuke .chan, Syntheyes, and 3D Equalizer are all supported, and adding additional export filters is rapid.)

The use of the XML format is very important, as it allows expansion of the amount and type of data being recorded by the take data file without breaking any existing pipelines that depend on reading the XML take data.

Stereoscopic Tracking and Metadata

Previzion now can read the metadata stream from the Element Technica and other specialized stereo rigs that send out IO and convergence data in real time. The optical accuracy of the lens calibration system makes it a natural fit for stereoscopic tracking, and the stereo tools in Previzion include the correct parameters to offset the measured pivots of the rig from the actual optical centers of the camera, which is critical to generate the actual IO and convergence points of the actual camera lens as opposed to the rig pivots.

Selecting a stereo rig in Previzion automatically causes the stereo metadata to be written into the take file, and the view from the left and right eyes can be switched to see how the CG and live action images match when the rig convergence and IO are modified.

Technical Specifications

Inertial/Optical Tracking: +/- 0.01-0.02 m positional accuracy, +/- 0.01 degree angular accuracy

Encoded Technocrane Tracking: +/- 0.02 m positional accuracy, +/- 0.001 degree angular accuracy depending on pan/tilt encoders [Kuper protocol presently supported, others upcoming]

GPS/IMU Tracking: +/- 0.02-0.05 m positional accuracy, +/- 0.1 degree angular accuracy

Lens Calibration: +/- 0.1 mm focal length accuracy at wide angles, +/- 5mm focal length accuracy at telephoto

Lens Controller Compatibility: Preston, RT Motion, Fujinon, Canon [C Motion under development]

Stereoscopic Rig Compatibility: Element Technica [Pace/3ality under development]

Live Video I/O: 2 x 4:2:2 Rec. 709 HDSDI inputs, 2 x 4:2:2 Rec. 709 HDSDI outputs

Video Playback: Up to 8 1920x1080 HDSDI MPEG2 files decoding simultaneously, depending on video and scene complexity and video frame rate

Interlaced Video Mode: Processes incoming 59.94i video and interlaced MPEG2 video at 59.94 frames/second to preserve motion signature

3D Rendering: 500k - 1M polygons, up to 8k of texture space, CgFx shading language, multiple video inserts, real time CgFx parameter adjustment and control

Keying, Color Correction, and Compositing: All operations performed in 32 bit floating point mode to avoid banding and artifacts in final image

Animation Transforms: 300-800 transforms/frame, or about 3 hero characters + simpler background characters

Data Logging Format: Collada XML, with camera data saved in matrix format as per Collada 1.4 spec.

Take Converter Output Formats: Maya ASCII, Nuke .chan, 3D Equalizer, Syntheyes [others added easily]

Future Work

Previzion's programmable architecture and open structure is designed to be very extensible. Our development will continue in the direction of more accuracy, photorealism, and speed. These areas include:

- 4:4:4 log camera input, keying, and output
- Linear, scene referred compositing and color correction
- HDRI illumination and 3D rendering
- Real time color grading with network connection to external control surfaces, such as Iridas
- Real time stereoscopic rendering using two Previzion systems in sync
- Real time facial animation rendering
- Programmable framebuffers, for grain and other 2D FX additions
- Real time greenscreen background lighting variation removal
- Embedding tracking data into HDSDI output frames

Production Usage and Results

On stage production use has driven Previzion's development from the beginning, and the feedback we have accumulated over several years of shows has shaped everything from the user interface to the format of the output files. Some of the productions and customers we have worked with include:

Knight Rider (2009 TV Series) Stargate Digital. Previzion was used for on set previsualization of driving shots in 'Knight Rider'.

Alice in Wonderland (2010 feature film) Sony Imageworks. 'Alice in Wonderland' was the first use of Previzion on a major motion picture. The system was used to track handheld and Steadicam shots through the course of principal photography, and the tracking data was cleaned up and used for several of the final shots in the finished Tim Burton film.

V (2009-2011 TV series) Zoic Studios. 'V' showed how Previzion could transform episodic television production. The system was used for on-set visualization of the interiors of the alien mothership, which existed only in CG, as well as for camera data recording. The use of the recorded data to drive a partially automated post rendering pipeline transformed studio internal production workflow, and enabled up to 400+ VFX shots per episode, which was unprecedented in episodic television production.

Pair of Kings (2010-2011 TV series) It's a Laugh Productions. 'Pair of Kings', a Disney children's TV show, used Previzion both for finished in camera composites, and tracking data capture for shots that were too difficult to finish in camera. Two systems were used to provide multiple camera angles.

Taxi Dog (2010 TV pilot) 'Taxi Dog', a children's television show, used Previzion to finish most of its shots in camera. The CG taxi and the photorealistic moving background provided a look that was very unique to children's television.

Dancing With the Stars (2010 TV opening sequence) CBS Digital. Previzion was used for the on set visualization and camera tracking data capture of the opening sequence for 'Dancing with the Stars'.

The Bold and the Beautiful (2011 TV opening sequence) CBS Digital. Previzion was used for the on set visualization and camera tracking data capture of the opening sequence for 'The Bold and the Beautiful.'

Tapas e Beijos (2011 telenovela) TV Globo Brazil. 'Tapas e Beijos' (Slaps and Kisses), a Brazilian telenovela shooting in 2011, uses Previzion to finish in camera shots looking out the windows.

Unnamed TV Pilot (2011 TV pilot) Stargate Digital. Previzion was the on set visualization system for a very high profile TV pilot that required extensive VFX to recreate the 1960s in multiple cities.

Unnamed 3D Feature Film (201x feature film) Previzion is the core on set tracking and visualization system of a major upcoming stereoscopic 3D feature film with extensive outdoor location shots involving animated characters.