Compositing Light Field Video Using Multiplane Images

Matthew DuVall, John Flynn, Michael Broxton, Paul Debevec

Google AI

thrive SIGGRAPH2019 LOS ANGELES · 28 JULY – 1 AUGUST

Problem

Light Field data sets require custom capture, processing and rendering techniques. Working with or modifying this data is even further complicated with the need for operations to be implemented within the rendering technology.

Multiplane Images

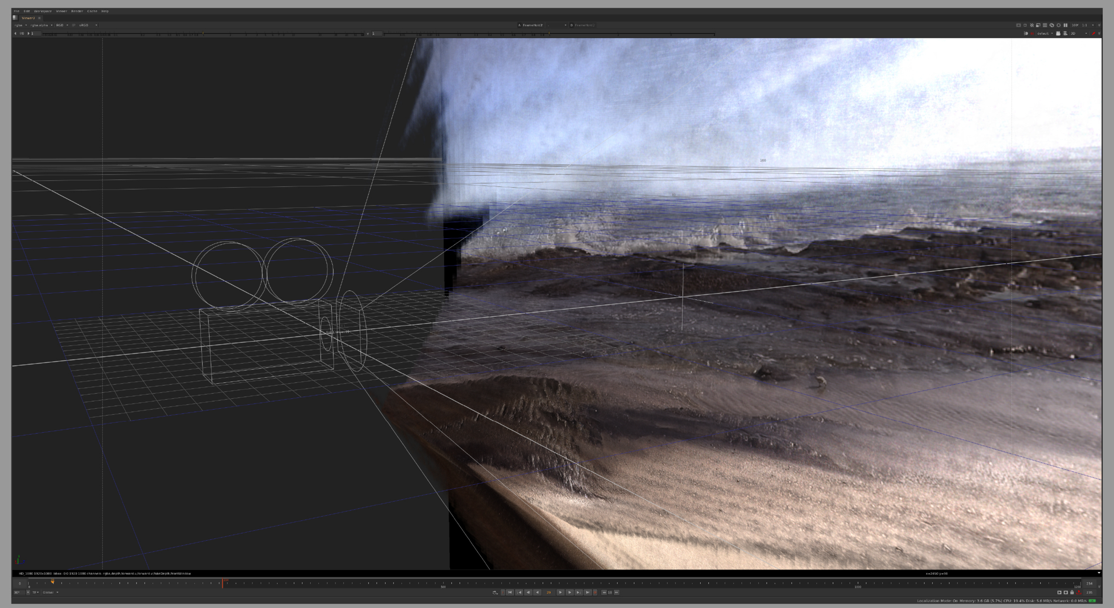
Multiplane images [1] encode three dimensional scene information as a series of front-parallel textured planes that fill a viewing frustum projected into a 3D environment. With an MPI it is trivial to render novel viewpoints simply by moving the virtual camera viewpoint relative to these planes. As such, MPIs can be utilized in most 3D rendering engines. Each plane not only displays a slice of the diffuse volumetric image, but also encodes view dependent effects such as reflections through the use of alpha compositing in back to front rendering.

a. A render of a multiplane image (MPI) from outside the viewing volume.

Workflow

To generate an MPI video, we capture image sequences of photography using a light field camera array. The array in our examples is 16 GoPro Hero 4 cameras capturing at a resolution of 2704 x 2028, at 29.97fps. The resulting mpeg4 data is processed through the DeepView solver [2] to MPI. We note that a MPI derived using DeepView shares several of the advantages of computer generated deep image composites [3]. with the streamlined abilities within Nuke that are possible when using DeepEXRs [4].

In this work we propose a Nuke based workflow (figure b.) and show examples that exploit the unique features of MPIs for generating 2D composites of 3D volumetric scene data.



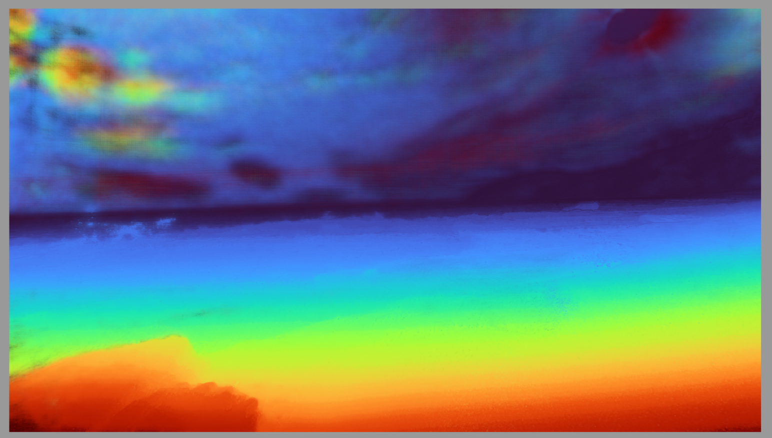
b. An MPI in Nuke's 3D scene environment.

References

[1] John Flynn, Graham Fyffe, Tinghui Zhou, Richard Tucker and Noah Snavely. 2018. Stereo Magnification: Learning view synthesis using multiplane images. Siggraph 2018.
[2] John Flynn, Michael Broxton, Paul Debevec, Matthew DuVall, Graham Fyffe, Ryan Overbeck, Noah Snavely, Richard Tucker. 2019. DeepView synthesis with learned gradient descent. *CVPR 2019*.
[3] Florian Kainz. 2013. Interpreting OpenEXR Deep Pixels.
[4] Tom Lokovik and Eric Veach. 2000. Deep Shadow Maps.

c. Stablized, reframed and defocused result

d. Center camera of the Light Field camera array



e. Composite depth derived from the MPI

Virtual Camera Editing

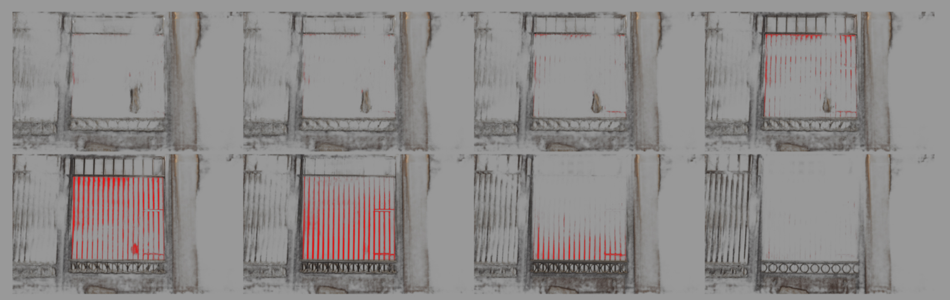
MPIs effectively encode a light field representation that supports high quality view synthesis at render-time.

This permits us to apply motion correction to achieve a smooth camera move when working with hand held footage. Rendering from a new, motion-stabilized camera trajectory while within the MPI viewing volume generates imagery that appears to have been captured with a steady-cam.

Another advantage of MPIs is that they can be processed to produce a view-dependent "composite" depth map (figure e.). The composite depth map is suitable for effects such as depth of field simulation (figure c.). The ability to display view dependent effects such as reflections and refractions translates through this composite depth map by describing the distance from camera to visible feature, rather than camera to geometry surface. As such surfaces such as water can be rendered more accurately with these effects than with traditional geometry depth maps.

Segmentation

To segment an object in an MPI, we must isolate the portions of planes containing that object in the RGBA textures. Restricting the matte projection within a specified range of depth representing the object and its desired surface effects completes the segmentation.



f.. Shown in red, the isolated portions of planes to remove.

g. MPI composite of all pixels on all planes (with virtual camera edits)

h. MPI composite with selective pixels disabled (segmentation and removal)

i. DeepEXR composite

j. Plate B.

k. Plate A.

Deep Compositing

Rather than segmenting portions of the image, the capture as a whole can be merged with another (figure i.). In our Nuke 3D environment, this is simplified to merging MPI scenes before rendering. In our video example "DeepEXR Compositing," the foreground MPI (figure A.) and background MPI (figure g.) were rendered to DeepEXR then merged.

Conclusion

We have presented a new method for compositing and manipulating video sequences in post-production using multiplane images in a popular commercial compositing tool, Nuke. We showed that MPIs generated with the DeepView solver can be used to achieve complex volumetric compositing and view synthesis effects without requiring substantial time and attention of a post production artist. Our approach shows the promise of light field video manipulation using MPIs for enhancing such post-production workflows.