Creating a Self-Contained Interactive Experience using USDZ

Scott Geffert; The Metropolitan Museum of Art; New York, NY/USA

Abstract

In the early phases of the pandemic lockdown, our team was eager to share our collection in new ways. Using an existing 3D asset and advancements in AR technology we were able to augment a 3D model of a collection object with the voice of a curator to add context and value. This experience leveraged the unique capabilities of the open Pixar USD format USDZ extension. This paper documents the workflow behind creating an AR experience as well as other applications of the USD/USDZ format for cultural heritage applications. This paper will also provide valuable information about developments, limitations and misconceptions between WebXR glTF and USDZ.

Background

The demand for the creation and delivery of public-facing interactive 3D and AR cultural heritage content is growing rapidly and has only accelerated as museum visitors have been physically separated from collections due to the Pandemic. The Metropolitan Museum of Art closed its doors to the public on March 13th 2020. While on full lockdown, our Imaging team set out to create and share a high-quality immersive AR experience via our website leveraging existing content and accessible tools. While we were not able to access the collection to digitize objects, our team did have access to our servers. Prior to the lockdown we had created a very successful 3D Photogrammetry model of a Zemi Cohoba Stand (Accession Number 1979.206.380). This is a signature collection object from the Arte del mar Artistic Exchange in the Caribbean exhibition that opened December 2019. The original purpose of the 3D imaging was to enable the creation of a digitally printed life-sized recontextualization by artist Miguel Luciano as part of his Met Civic Practice Partnership Artist in Residence.

One of the great joys of being part of the cultural heritage imaging community is learning about the objects from our curators and conservators as we engage in our work. During the imaging session we learned of the history and important ceremonial role of this object and felt that it would be valuable to convey this behind-the-scenes experience via AR technology. In 2019 we had been exploring the benefits of the Pixar USD and USDZ extension and experiments using Adobe™ Project Aero™ and the Apple™ AR Quickview™ and Reality Composer™ tools. In the course of our explorations we tested the ability to add audio to 3D objects as part of Augmented Reality tests. The driving force was to recreate the educational experience of a "curator over your shoulder", or a "curator in your living room". The combination of AR and Audio was very powerful, but we were concerned that the only way to embed audio into a self-contained AR experience was to save the asset in the Apple .reality format; a proprietary format specific to a single (free) application. Fortunately, in late March 2020 we discovered an almost casual mention buried in an Apple developer technical update regarding the ability to embed audio into the open USDZ format. This single update enabled the ability to share an entire AR experience including bilingual audio and user interface elements together with the high-quality 3D model to our global audience of iOS™ and MacOS™ users. To our knowledge, this was the first ever use of this unique self-contained USDZ capability for cultural heritage AR. With this background and context, we will outline the technical workflow beginning with the original 3D model to the delivery of the final experience. The general workflow is available on a dedicated blog page, augmented-reality-zemi-arte-del-mar but we will provide greater detail below.

Workflow: From Model to Interactive AR

3D Model Creation/Preparation

Our imaging process started with a capture session in the fall of 2019 that included Image Production Associates Chris Heins, Heather Johnson, and our Deputy Chief Photographer Oi-Cheong Lee. It was a follow up to their recent photogrammetry training. The 3D model processed in Metashape™ was quite successful and captured the nuances of this carved wooden object. To improve the surface qualities of the texture, the roughness map was edited by Imaging Production Manager, Wilson Santiago, using Adobe Substance Painter™. The gloss value of the inlaid shell teeth was adjusted along with other matte/gloss areas of the carved wood.

3D Model Optimization

Creating a high-quality 3D model can be a challenging endeavor, but optimizing a high-poly 3D asset for online delivery is a difficult workflow to sustain across a workgroup due to an almost complete lack of standardization coupled with the unique operator skills necessary to retopologize, unwrap, and bake down a model. This process is highly subjective and time consuming. In the case of this particular project, we had an opportunity to explore two approaches: manual optimization and automated optimization. From an artistic point of view, one could argue that a manually optimized 3D asset could provide the highest quality, but from a sustainable mature distributed workflow point of view, manual optimization can be problematic. For the Wood Deity Figure, Wilson Santiago manually optimized the original 20 Million Polygon scan to a compact 4500 Polygon model while preserving the high-quality texture detail. This manually optimized model provided a comparison to gauge the quality of any automated optimization tests.

For automated optimization we turned to InstaLOD™ software https://instalod.com/. InstaLOD is a toolkit for mesh processing that includes automating mesh optimization, UV unwrapping, conforming normals and baking. This solution creates finished optimized files, but unlike other automated tools, the configurable workflow allows our team to optimize the process for our geometrically complex artworks. Working with the InstaLOD support team we arrived at an optimization quality that rivals the manual approach in a fraction of the time.

The original model is loaded into InstaLOD and the user selects an appropriate custom preset. In this specific case, the preset is a processing pipeline that begins with retopologizing to a target of 100k polygons, UV unwrapping and baking textures to a specific 2K resolution. As part of the InstaLOD process additional maps are generated including Metalness, Occlusion and Roughness. The software supports various output formats but we find that .obj or .fbx work best for the downstream conversion workflow.
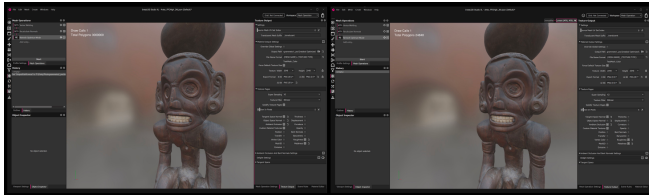
InstaLOD: 30 Million Polygon model (left) Automatically optimized to 24000 Polygons with 2K texture baked (right) in under 3 minutes

USDZ Conversion
The easiest and most reliable method for USDZ conversion is to use the free Apple Reality Converter™ software. This software is available from the Apple Developer program web site: https://developer.Apple.com. The simple drag and drop interface tools allow you to visualize the impact of the individual texture maps. There are fields for embedded metadata and attributes such as scale as well. One of the strongest features of Reality Converter is the ability to immediately share the file to your iPhone™ or iPad™ via Airdrop™. Due to a lack of standardized base measurement units in 3D imaging content creation tools, it is often the case that a model will present in AR at the wrong size. This AR preview is invaluable to verify the embedded units are correct. It's also extremely valuable to view the final product on a target device before Exporting the USDZ asset.
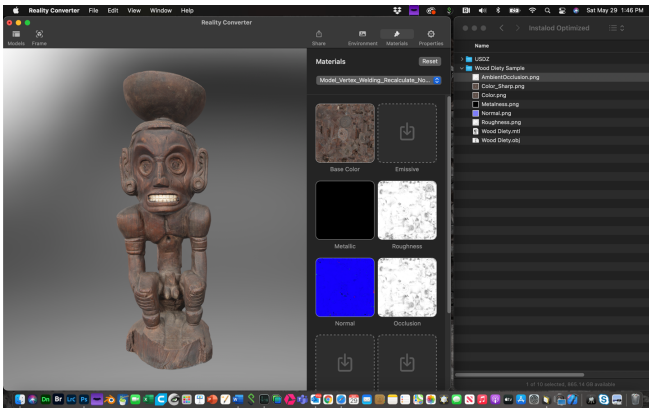

Reality Converter: Drag and Drop USDZ creation and validation with immediate AR preview

Bringing it all together: Creating an interactive AR
With the optimized file converted to a validated and properly scaled USDZ asset we leveraged the Apple Reality Composer application to build the interactive experience. The best way to describe Reality

Composer is to think of it as Keynote™ for Augmented Reality. You drag elements into the working stage, apply behaviors, and play your "show". It's a very intuitive tool-especially on the iPad, but it has its limitations. Reality Composer is available for free from the iOS App Store, but it is also part of the MacOS XCode toolkit available through the Apple Developer Program. When using the MacOS version the program syncs interactively with your iOS devices which is helpful for making design tweaks.
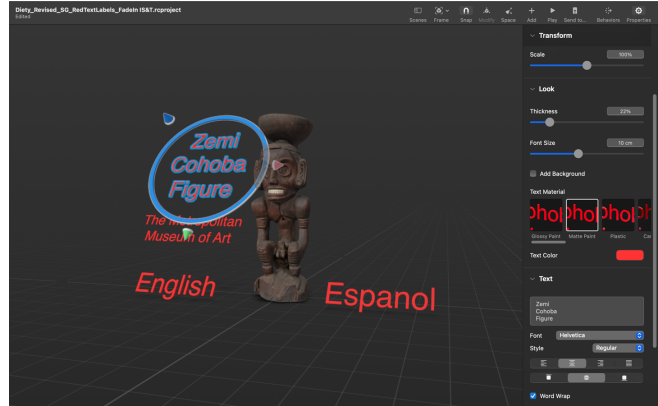
Reality Composer: Drag and Drop AR scene elements and assign attributes, actions and interactivity. Save directly to USDZ

When launching the application, you are prompted to select an anchoring option. A horizontal Anchor is the most common, but you are able to designate vertical planes (walls) faces, images or even a custom scan of an object. With the horizontal plane selected, you then add the 3D USDZ asset to the scene. Next text is added. In our case, we decided to incorporate English and Spanish curator narration and needed to prompt users with the options available while keeping the interface as simple as possible. Assistant Curator for the Art of Ancient Americas Department James Doyle recorded the audio on his phone from his home during lockdown. Nina Diamond from our Digital Department provided guidance on refining the script and Paul Caro of our Digital Department helped trim and clean up the audio also working from home. With the buttons roughed out it was time to add the audio interactivity. After selecting the English text box, the English audio track was attached to the button. The same process was repeated for the Spanish audio track. With the audio tracks wired to the buttons, the next step was to iterate testing various timings for fade-in and fade out to give just enough information to the user. It's important to note that for a successful experience the user needs to feel they are in their own environment with a curator. Any menus, hot spots, or other distracting elements need to disappear to keep the focus on the art.

Iterative testing
Once we arrived at the desired baseline functionality, we saved the project using the USDZ option from Reality Composer. This writes everything necessary down into the self-contained open USDZ format. The USDZ was shared via iMessage across the team and we discussed fine tuning. Madhav Tankha from our Digital Department joined the collaborative testing team. The greatest challenges were related to the selection and design of the text elements. When

placing text for AR there are some necessary considerations. Type cannot be very small, and no matter what color or style you choose you will find that the text will disappear against the background depending on where the user is viewing the experience. There were other changes such as angling the text for legibility. For example: if the user places the object on a coffee table, sitting on a couch the user would have to bend down to easily read the type placed vertically. If the type was placed horizontally, the user would have to stand over the object to read the type. By placing the type at 15-degree angle users were able to easily read and interact with the text prompts from any vantage point. Another change was related to how people view AR objects. Some users immediately wanted to scale and rotate the object with their fingers, and others preferred to walk around the object as the curator spoke. A few users noted that they live in small apartments and were unable to walk all the way around the object. Since all of this feedback was valid, we decided to add a simple tap-to-rotate function. This solution functioned well, and would not have been identified without this important testing and feedback.

Publishing on the web

Barbara Bridgers (our former department head) and I worked with Xue Chen and Deepa Paulus from our 3D team on a creating a dedicated informative blog page with the help of William Fenstermaker from our Digital Department. Incorporating a USDZ file on a web page requires placing a few lines of HTML code to enable a Quicklook badge to inform users that an AR experience is available. It's important to note that interactive AR USDZ files can be shared directly via iMessage, email, Airdrop or file system. because they are fully self-contained there are no sidecar files or folder trees to manage. A USDZ file does not require HTML code to enable its interactivity. A number of people have reviewed our USDZ AR experience and some have incorrectly referred to it as WebXR based. It is not.

Zemí Cohoba Stand
A.D. 974–1020
Taíno

⌂ LISTEN

WebXR: An example of an internal test of a cross-platform Interactive AR WebXR object with Audio

At the time this experience was created, the toolkits for WebXR were still in Beta testing and the Google™ Model Viewer™ shipped in June 2020. As soon as this capability was available, we set out to explore building a WebXR version of the experience and have been successful linking audio and interactivity to extend this experience to an even wider cross-platform audience. Working with WebXR could likely be the topic of a future technical paper. Since we shared our original USDZ based experience we also worked to explore recreating the same experience as an Instagram™ filter using the Spark AR™ development platform. At this point in time, purely web-based AR experiences that do not leverage native Apple and Android™ AR viewers suffer from problems with anchoring (objects float off into space), environmental lighting, low texture detail, color fidelity and cluttered user interfaces.

USDZ and WEB XR

There are key conceptual and architectural distinctions to keep in mind as we explore use cases for the creation, application, and delivery of 3D AR content: USDZ is an extension of the open USD file format developed by Pixar™. USD was primarily designed as an interchange format for 3D content creation. USD is positioned as a "long game" format that was refined through in-house production at Pixar for over a decade before being open sourced in 2016. The strength of USD is the ability to store and communicate complex scenes comprised of many 3D elements across multiple users and applications. The adoption of USD has grown exponentially in the entertainment and architecture industries, and this has only accelerated since Apple and Adobe decided to adopt and help extend the format. Autodesk Maya™, Houdini™, Nvidia™, Blender™, Epic™ and Unity™ are embracing USD primarily due to its rich deep, modern feature set and the growing industry necessity for a common 3D interchange file format. On June 8th, 2021 3D file sharing site Sketchfab™ announced support for upload of USD files including USDZ, USDA, and USDC. This bodes well for USD as a potential future common archival format for cultural heritage 3D content. Archiving 3D data is a challenge facing cultural heritage users worldwide.

This video from Nvidia highlighting its USD based Omniverse™ collaboration platform is a good overview of the strength of USD https://youtu.be/vFxytzQlOEs. It is also valuable to note that Guido Quaroni, a 24-year veteran of Pixar software R&D has recently joined Adobe.
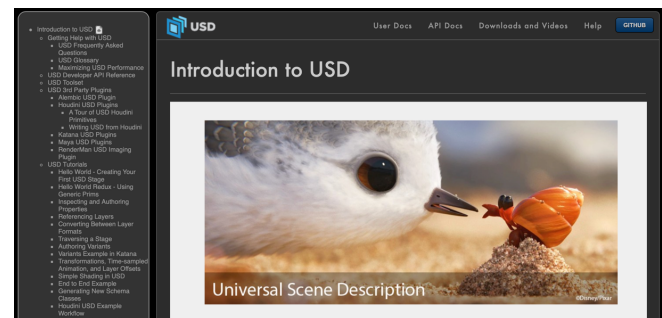
*Pixar USD: A modern open interchange format for 3D graphics production*
https://graphics.pixar.com/usd/docs/index.html

WebXR is a critically important development for content delivery and at least for now relies on the glTF transmission format promoted by the industry consortium The Kronos Group™. glTF is a transmission format promoted as a "3D JPEG" and feature by feature quite similar to the USDZ subset of USD. A critical distinction is that in the specific use-case of adding audio to a web-based AR experience, instead of using a user-friendly application such as Reality Composer™ to build the interactivity into the asset, the experience must be generated at the HTML code level. For WebXR-based experiences the 3D models (USDZ for iOS and glTF for Android™), text, audio tracks and any interactive logic need to be carefully programmed and debugged by a web developer. If your institution has the staff bandwidth to create custom code for each AR experience, this is a viable option. On May 19th 2021 Google updated its WebXR model viewer web component "to generate a USDZ on-the-fly when the user clicks the Enter AR button". While this seems like a harmless convenience feature, it actually encourages the use of glTF over native USDZ content development. Thankfully it is still possible to serve both USDZ and glTF files via Model Viewer to provide the best possible experience on Apple and Android devices. The current state of 3D and AR formats is unfortunately similar to the VHS/Betamax debacle in the late 1970's. On the positive side, WebXR eliminates the need for dedicated Apps and can help extend experiences across any platform, but to the average museum photographer, designer, curator, conservator or exhibition designer, the development skills required for WebXR are out of reach. It's the complexity of creating and sharing web-based 3D content that has led some cultural heritage users to rely on third-party services. The exciting aspect of USD and USDZ is the broad accessibility to mainstream content creation and interchange tools for cultural heritage creative professionals.

Practical Applications of USDZ

The ability to create and embed a complete AR experience into a single portable and open USDZ file has far reaching implications for cultural heritage. An example would be the ability to take multiple fragments of an object and drag them into a single interactive USDZ scene using Reality Composer or equivalent USDZ compatible DCC software tool. The USDZ document could be easily interchanged across different users and applications as the object positions are modified. Audio notes, text and even supporting photographs can be added and archived within a single USDZ document. Of course, this work can take place on the desktop with the help of a highly skilled operator using a complex 3D Augmented Reality development software program. It could also be accomplished with the aide of a web coder using WebXR, but there are many cultural heritage use cases that benefit from ease of use, accessibility and interchange inherent in the USD(Z) format and the growing number of companies and software applications that support it.
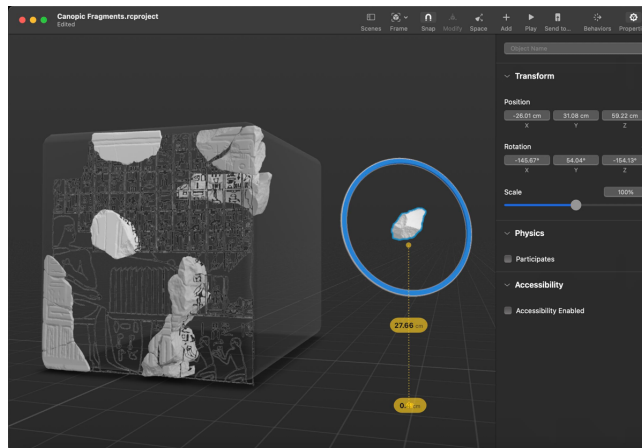
Ease of Interchange: In this example multiple USDZ models of Canopic Jar fragments are easily placed in a virtual reconstruction using the free Reality Composer™ App on an iPad™. The exported USDZ scene is easily shared and modified or annotated across multiple users with no programming or coding experience

Conclusions

The creation of this self-contained interactive AR experience was driven by a desire to share a part our collection during the height of the pandemic while our galleries were closed to the public. We leveraged a new technical capability launched only weeks prior with many unknowns, yet thousands of people around the world have enjoyed this single experience. The work helped establish an internal quality benchmark that has enabled us to compare quality and development workflow complexity across different AR delivery platforms (WebXR/Instagram Spark).

It is critically important to underscore that any device or operating system is fundamentally capable of reading this open format and that any incompatibility with the open USD or USDZ format is a result of competitive marketing and not due to core technical limitations. We will be watching developments closely while actively exploring WebXR, other 3D file formats and delivery mechanisms as the industry matures. In the meantime, broader adoption of USD (and USDZ) would certainly benefit the cultural heritage community and could play an important role in archiving 3D cultural heritage content.

Author Biography

Scott Geffert is General Manager of The Metropolitan Museum of Art Imaging Department. Responsible for overseeing 2D photography and 3D Imaging, color management, and the development and implementation of ISO imaging standards. His interests and involvement in imaging have evolved over the years from supporting users worldwide to taking an active role in helping steer the industry via advocacy of international standards and involvement in the IS&T, ISO and CIE organizations. Scott is currently heading up ISO standards work to define best practice in spectral imaging. Scott's long-term experience in color management especially within the cultural heritage community has led to innovative patented work in the field of multi-spectral LED lighting technology, enabling precision tunable illumination for viewing and digitizing artworks.