Editors: Cláudio T. Silva, csilva@nyu.edu | Daniel Weiskopf, weiskopf@visus.uni-stuttgart.de

A Flight through the Universe

Miguel A. Aragon-Calvo | University of California, Riverside
Mark SubbaRao | Adler Planetarium

In the early part of the last century, Edwin Hubble and others discovered that the *spiral nebulae* were galaxies in their own right, collections of hundreds of billions of stars far beyond our Milky Way. With this discovery, our universe became much, much bigger. Mapping these galaxies uncovered clusters and superclusters of galaxies, revealing that these structures fit into a larger network of filaments, walls, and voids—the cosmic web.

Tracing the cosmic web meant determining distances to hundreds of thousands of galaxies, something that wasn't possible until the beginning of this century. The Sloan Digital Sky Survey (SDSS) was the largest systematic cosmic cartography undertaking, astronomy's version of the genome project. The main survey lasted from 2000 to 2008, digitally imaging one-third of the sky in multiple color bands and discovering 200 million galaxies. Light spectra were taken and distances measured for approximately 1 million of these galaxies, allowing us to map out the cosmic web's 3D structure for the first time. For each galaxy in a "spectroscopic sample," we have imaging in several bands and an extensive list of properties, including color (defined in astronomy as the difference between the flux observed by two filters), concentration (the ratio between the total light inside 10 and 90 percent of the galaxy's radius), shape, distance, and so on. In addition to computer-generated parameters, we used a catalog of galaxy classifications from the Galaxy Zoo project (www.galaxyzoo.org), one of the most successful citizen science undertakings, with more than 150,000 volunteers who classified millions of SDSS galaxies by eye, a task where humans are still better than computer vision algorithms.

The SDSS continues to this day, with survey extensions whose various scientific goals include discovering supernovae, mapping the structure of the Milky Way, searching for exoplanets, and mapping out even more distant galaxies.

 1521-9615/15/$31.00 © 2015 IEEE Copublished by the IEEE CS and the AIP November/December 2015

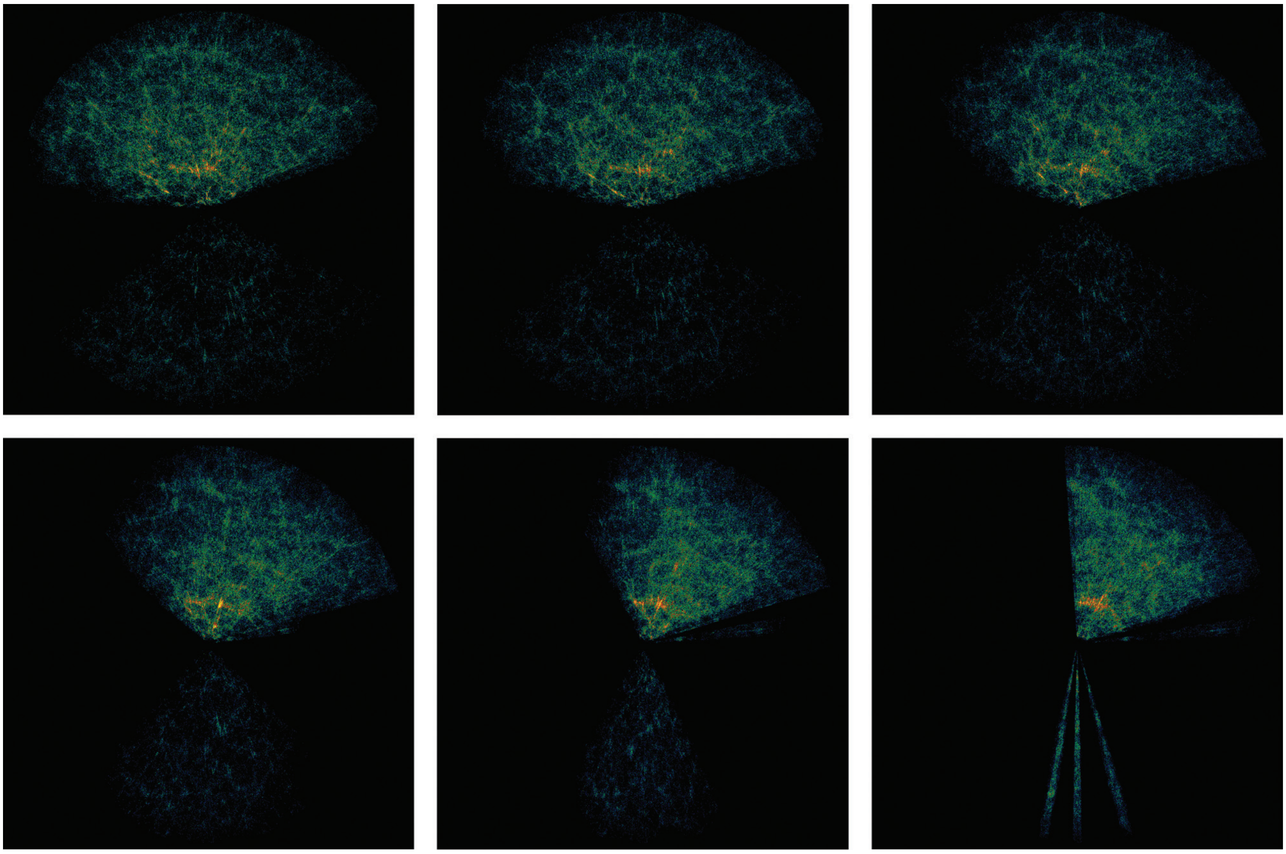

Figure 1. Survey geometry showing the complex shape characteristic of redshift surveys. The Earth is located at the wedge's apex. Galaxies are represented by dots of light.

In this installment of Visualization Corner, we'll take a look at the creation of a fly-through animation across the largest map of the universe to date, and how we addressed the scientific, aesthetic, and technical challenges posed by this project.

Scientific and Artistic Goals

With our visualization, we wanted to give the viewer the experience of flying through the largest structures in the universe, viewing galaxies as islands of light in the vast ocean of the cosmic web. We attempted to respect the sheer immensity of space: the pacing of the visualization, the flight-paths, and the rendering of the galaxies were all done in a way to make the distances seem huge.

A key creative tension exists between making the visualization seem immersive and illustrating the universe's large-scale structure. Immersion comes from passing close by a galaxy and then watching it fade away into nothingness as we fly away. The large-scale structure is best shown when we view the galaxies at all distances and aren't distracted by large ones in the foreground. We sought to strike a balance between these two goals, providing a cinematographic fly-through of foreground objects, while rendering objects in a way that the light from distant background galaxies blends together into a map of the cosmic web.

We also attempted to accurately represent the relative sizes and types of galaxies. Many of the more distant galaxies have poorly resolved images—by representing them next to similar, closer, better resolved galaxies, we get a homogeneous view of the survey at all distances. Doing so allows the visualization to show how the mix of galaxy types varies with environment, the densest clusters being dominated by giant elliptical galaxies, and the spirals avoiding the densest regions.

We designed the visualization to minimize the appearance of artifacts coming from survey design or measurement systematics, allowing the viewer to focus on the distribution of galaxies and the structure of the cosmic web. The first thing you notice when plotting positions of SDSS galaxies are the "watermelon slice" structures (see Figure 1) resulting from

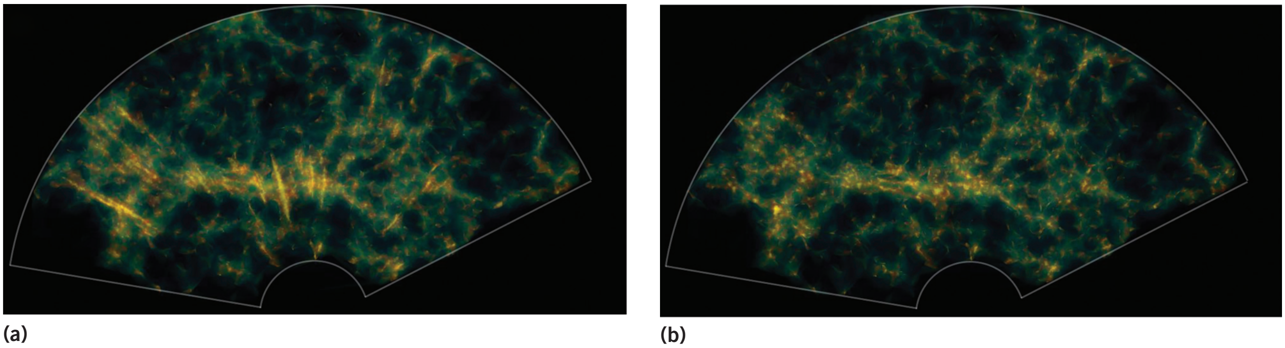

Figure 2. Fingers of God: (a) original distribution of galaxies, showing fingers of God artifacts as lines pointing to Earth, located at the cone's apex, and (b) galaxy distribution after correction for fingers of God. Not only are there fewer distracting features, but the cosmic structures seem sharper.

the survey's footprint on the sky. We designed the visualization to avoid showing the survey's boundaries and instead focus on the structure within the slices.

Another artifact is the way in which galaxy distances are measured. From the galaxy's spectrum, we measure a redshift to see how much the universe expanded while the galaxy's light was traveling to us. The motion of the galaxy along the line of sight also contributes to redshift, producing an error in the measured distance. In massive clusters where the galaxies are moving with very fast random trajectories, these errors lead to linear structures along the line of sight (because we can only measure the velocity vector's radial component) also known as "fingers of God." These are no real physical structures, and because fingers of God point toward Earth, they can become very distracting. Although we have no way of correcting this distance error for a single galaxy, we can use statistical techniques to identify and collapse these fingers of God, leading to a more accurate visualization. We identified the fingers of God by selecting galaxy groups with high elongation and aligned along the line of sight. Each finger of God was then "compressed" in the radial direction to make it isotropic (see Figure 2).

Scientific and Technical Challenges

During the animation's design, we faced several challenges in trying to combine the accuracy required by a scientific visualization with the artistic license needed to produce a compelling experience. The physical scales covered in this galaxy map are beyond our normal experience by many orders of magnitude. Even the speed of light is insignificant at these scales—it would take light more than 1 billion years to travel across the map. Distances between galaxies are so vast compared to their sizes that galaxies look like little speckles of light in the vast emptiness of intergalactic space.

In addition to this, light dilution reduces the apparent brightness of distant galaxies, producing a large dynamic range in brightness between near-by and far galaxies. The main technical challenge we faced was the rendering of 400,000 individual galaxies with different sizes, shapes, and color and with a high level of detail. Each of the 400,000 galaxies has a rich internal structure consisting of spiral arms, star-forming regions, bars, and so on. Due to limitations imposed by the telescope and the atmosphere, most distant galaxies are poorly resolved and can't be used for high-resolution rendering.

Mocking Up 400,000 Galaxies

Galaxies have a wide variety of colors, shapes, morphologies, and sizes. Color and shape are good indicators of a galaxy's history. As such, we expect that galaxies with similar shapes also have, generally speaking, similar colors, and vice versa. Even though individual galaxies have unique properties, they can be roughly classified, based on their morphology, into two main categories, or Hubble types: ellipticals and spirals. Ellipticals, consisting mostly of old galaxies, are red and concentrate their light in their center, whereas spirals actively produce new stars, are blue, and have a light distribution that's much more uniform. This classification, widely used in astronomy, lets us compress the global properties of galaxies with one single value. We extended this idea to include not only morphology but also color (defined as the difference between the total light flux in the g and r filters, roughly corresponding to green and red) and light concentration (how concentrated the light is at the galaxy's center). Using these parameters, we classified galaxies into 256 types covering the full range of observed morphology colors and concentrations. Each of these types could then be used to represent each of the 400,000 individual galaxies in the survey.

Next, we created a template of 256 galaxies with high-quality images, providing a uniform coverage of the color-concentration parameter space. To do this, we started by computing the probability density function (PDF) of galaxies in the color-concentration parameter space (see Figure 3) from the full sample of galaxies by using a Delaunay-based interpolation scheme. We then populated this PDF with galaxies via a Monte Carlo scheme and 256 samplings. The sampling points received a more uniform distribution by using a multistep Voronoi centroid algorithm.

For each sampling point, we found the closest galaxy in the parameter space, considering only those galaxies with the following criteria: spiral face-on galaxies with high-quality imaging (more than 256 pixels of diameter, diameter defined as the ellipse containing 90 percent of the light), no close or interacting companions within the target galaxy's radius, and no bright stars in the foreground or large galaxies in the background inside the galaxy's radius. Our template thus only contains spiral galaxies because ellipticals can be easily simulated using sprites with a radial light distribution. We further de-projected each galaxy in this template to make it appear as close to face-on as possible, with a circular mask applied by using automated scripts. Finally, we visually inspected each galaxy and removed any remaining stars by hand.

We assigned each of the 400,000 galaxies to its closest template galaxy in the color-concentration parameter space. While we did template matching for all galaxies, in practice, we only used the subsample of spiral galaxies because ellipticals can be closely approximated with camera-facing billboards (sprites); Figure 4 shows several examples of spiral galaxies included in the final template.

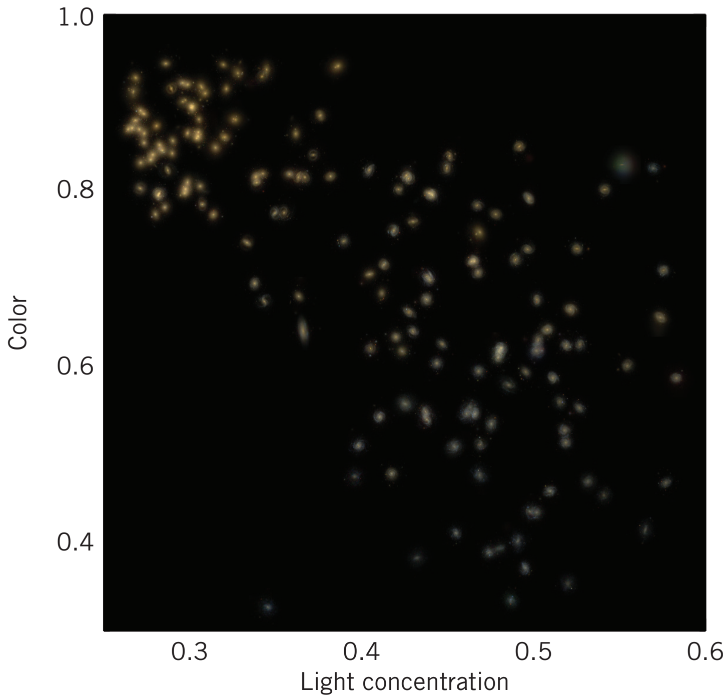

Figure 3. Distribution of galaxies in the concentration-color space. Elliptical galaxies are red and highly concentrated, whereas spirals are blue with a more uniform light distribution.

Rendering

We did the 3D scene construction and rendering by using the popular open source Blender (www.blender.org) software, which a large community of artists and an increasing number of scientists use for 3D modeling. The main deciding factor in using Blender was its Python API, which allows control of nearly all of its modeling and rendering features. Being based in Python meant it required minimal effort on our part to write a script to construct the 3D scene. Figure 5 shows a code snippet with the functions to load galaxy images and assign them to Blender materials. The Python API is straightforward and allows us to control properties of the materials such as alpha value, specularity, emissivity, and so on.

Template Bins

Because we mapped all the galaxies to a template of 256 galaxies, there are only 256 unique images. We then created one mesh object in Blender for each of the 256 images. Each Blender object contained the corresponding galaxy image and the positions of all the galaxies assigned that image. We divided the template into elliptical and spiral galaxies, and treated each morphological component differently.

For spiral galaxies, we generated a quad with a random 3D orientation and size scaled with its physical size (see Figure 6). The space between galaxies is very large compared to their physical size: if we were to display them at their actual size, we would only be able to see the most nearby and largest structures; the rest of the galaxies would look

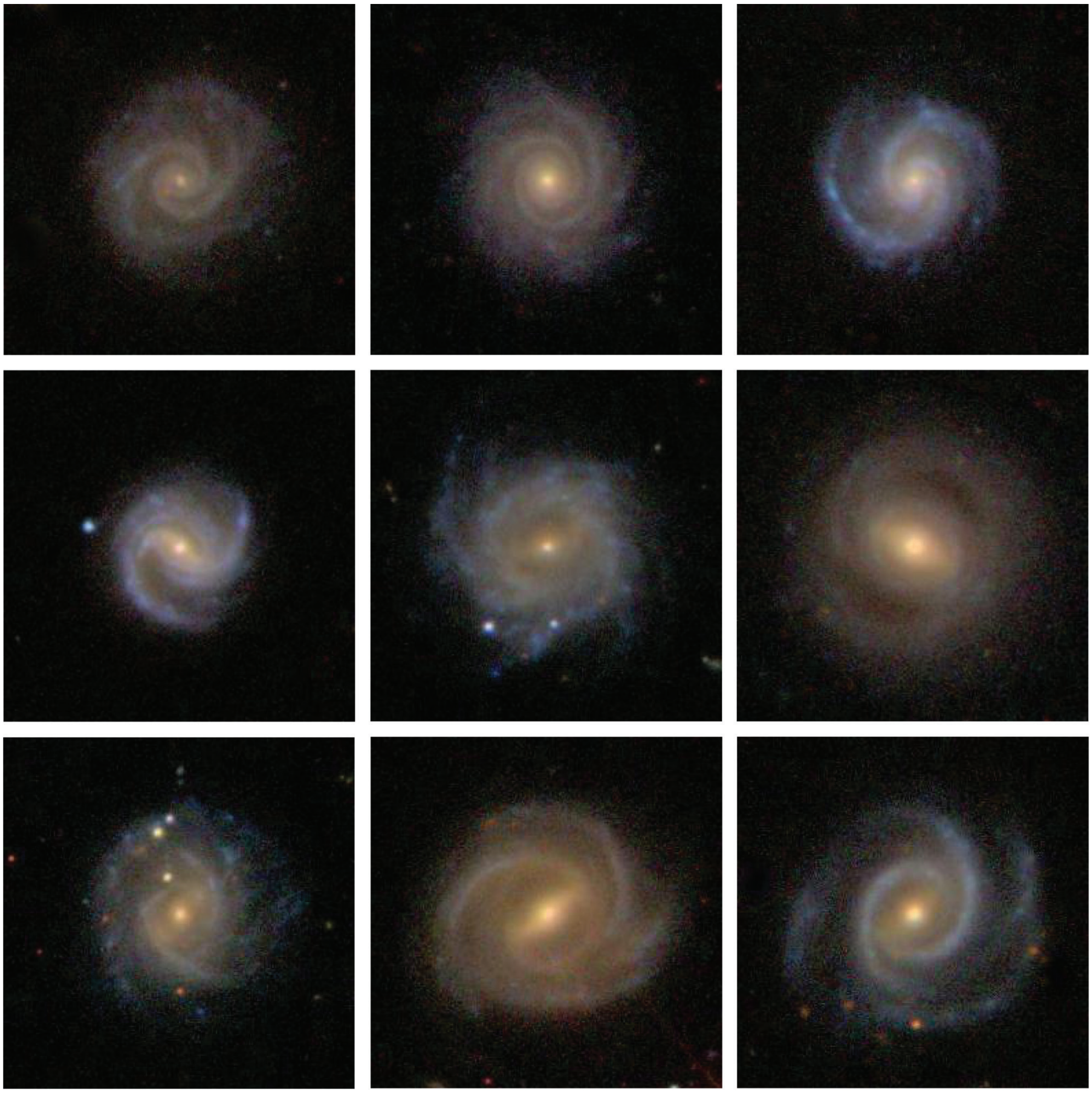

Figure 4. Template of high-resolution galaxy images used in the SDSS animation.

like unresolved stars. We added a scaling factor of 10 to all galaxies to compensate for this, choosing this empirical factor to resolve most galaxies within the typical scale of the cosmic web and avoid overcrowded-looking regions.

Elliptical galaxies, in contrast to spirals, have a tridimensional shape and can't be rendered as an oriented quad, as mostly flat spiral galaxies are. A screen-oriented quad, also known as a billboard, is a good solution for spherical objects, but for ellipsoids, it always presents the exact same orientation to the camera, producing an undesired rotation effect. We instead simulated simple tridimensional ellipsoids by placing three billboards along a line corresponding to the ellipsoid's major semi-axis. The line of sprites was randomly aligned in 3D and scaled with the galaxy's physical size; we made the central billboard larger than the adjacent billboards to better fit an ellipsoid.

Camera Setup and Path

We rendered the scene in 3D stereo by using a fixed camera rig with cameras pointing along parallel lines. The rendering was done with supersampling to reduce aliasing artifacts; we rendered images for each camera to a slightly wider format than 1080p to allow room for later adjustments of the 3D effect.

We designed the camera path to give a good view of the vastness of intergalactic space and, at

```
#--- Load galaxy image into texture
def  load_texture(filename,GalID):
     # Blender Texture
     tex = Blender.Texture.New('GalTex'+str(GalID))
     tex.setType('Image')
     tex.setImage(img)
     tex.extend=Texture.ExtendModes.EXTEND
     tex.imageFlags=Texture.ImageFlags.INTERPOL |
                    Texture.ImageFlags.USEALPHA |
                    Texture.ImageFlags.MIPMAP
     #--- Material
     mat = Blender.Material.New('GalMat'+str(GalID))
     mat.setAlpha(0)
     mat.emit = 1
     mat.setSpec(0)
     mat.mode |= Material.Modes.ZTRANSP
     mat.mode |= Material.Modes.NOMIST
     #--- Add texture

mat.setTexture(0,tex,Texture.TexCo.UV,Texture.MapTo.COL |
Texture.MapTo.ALPHA)
     return mat

mats = []
for i in range(0,255):

     mats.append(load_texture('gal-'+'%(#)03d' %{"#": i}+'.png', i))
```

Figure 5. Blender python API. We show the code that loads the galaxy template images into Blender materials. The API provides full access to Blender functionality, making it perfect for automated scripts.

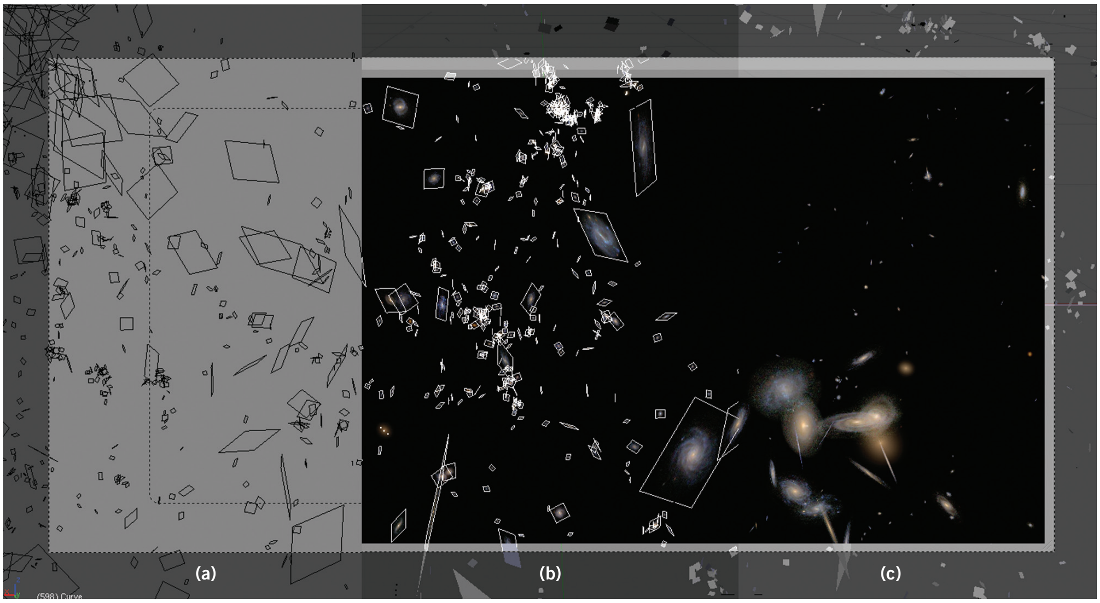

Figure 6. Composite view of (a) the wireframe scene, (b) the textured polygons, and (c) the final rendering. The space between galaxies is very large compared to their physical size: if we were to display them at their actual size, we would only be able to see the most nearby and largest structures; the rest of the galaxies would look like unresolved stars.

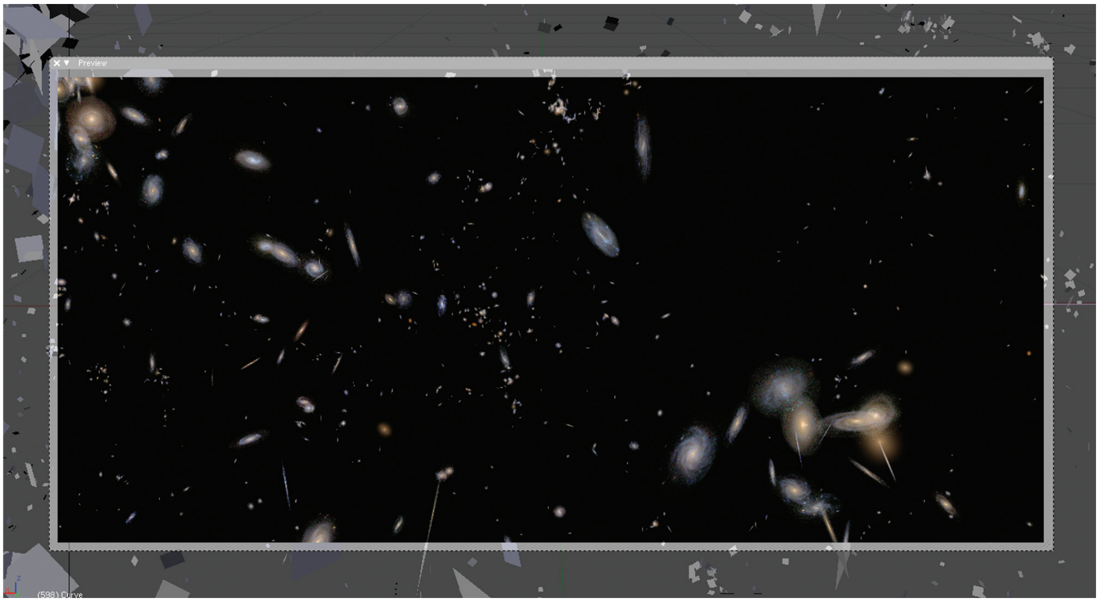

Figure 7. Final rendering used in the fly-through. The image shows both the detail in nearby galaxies as well as the large-scale structures they delineate. Distant galaxies are dimmed to reduce background distraction.

the same time, show several astronomically interesting regions, as we intended to use this animation for scientific talks. We manually made the camera path to include a flight through a large cluster and several filaments, and transversal pans across large voids. The camera path makes a close loop. We avoided eye-pocking galaxies and only show one close pass near a spiral at the beginning of the loop.

We did the rendering on a relatively small workstation with eight cores and 32 Gbytes of RAM; it took half a day to render the complete animation.

Figure 7 shows a snapshot from the final rendering, with two groups of galaxies and a large void in the background. The galaxies in the distant background are located on the far side of the void and delineate other groups and filaments.

The SDSS fly-through was a very rewarding project. As we fly in between cosmic structures, we can begin to grasp the overwhelming magnitude of the universe and our very small size in the order of things. Even traveling at the speed of light, it would take us more than 2 million years to travel to the nearest spiral galaxy. Effectively, the vast majority of the universe is out of reach. We're stranded in our own "island universe," but it doesn't prevent us from exploring faraway regions using powerful telescopes in our quest to understand how the universe works.

The SDSS fly-through animation has been extensively used in scientific talks, documentaries, and 3D concerts. In the hope that the galaxy map presented here will be useful for other people, we've set up a dedicated webpage to host galaxy positions, the galaxy image template, Python scripts, and blender files. High-quality renderings in mono and stereo are also available for download—visit www.dataviz.science/HOME/sdss_fly.html for more.

Miguel A. Aragon-Calvo is a visiting research scientist at the University of California, Riverside. His research interests include galaxy formation, the large-scale structure of the universe, data visualization, and big data. Contact him at maragon@ucr.edu.

Mark SubbaRao is director of the Space Visualization Laboratory at Adler Planetarium. His research interests include cosmology, scientific visualization, and the Large Synoptic Survey Telescope project. Contact him at msubbarao@adlerplanetarium.org.